Noncommutative Geometry, Quantum Fields and Motives

Alain Connes

Matilde Marcolli

A. Connes: Collège de France, 3, rue d'Ulm, Paris, F-75005 France, I.H.E.S. and Vanderbilt University
E-mail address: alain@connes.org

M. Marcolli: Max–Planck Institut für Mathematik, Vivatsgasse 7, Bonn, D-53111 Germany
E-mail address: marcolli@mpim-bonn.mpg.de

Contents

Preface

The unifying theme, which the reader will encounter in different guises throughout the book, is the interplay between noncommutative geometry and number theory, the latter especially in its manifestation through the theory of motives. For us, this interwoven texture of noncommutative spaces and motives will become a tool in the exploration of two spaces, whose role is central to many developments of modern mathematics and physics:

- Space-time
- The set of prime numbers

One may be tempted to think that, looking from the vantage point of those who sit atop the vast edifice of our accumulated knowledge of such topics as space and numbers, we ought to know a great deal about these two spaces. However, there are two fundamental problems whose difficulty is a clear reminder of our limited knowledge, and whose solution would require a more sophisticated understanding than the one currently within our immediate grasp:

- The construction of a theory of quantum gravity (QG)
- The Riemann hypothesis (RH)

The purpose of this book is to explain the relevance of noncommutative geometry (NCG) in dealing with these two problems. Quite surprisingly, in so doing we shall discover that there are deep analogies between these two problems which, if properly exploited, are likely to enhance our grasp of both of them.

Although the book is perhaps more aimed at mathematicians than at physicists, or perhaps precisely for that reason, we choose to begin our account in Chapter 1 squarely on the physics side. The chapter is dedicated to discussing two main topics:

- Renormalization
- The Standard Model of high energy physics

We try to introduce the material as much as possible in a self-contained way, taking into consideration the fact that a significant number of mathematicians do not necessarily have quantum field theory and particle physics as part of their cultural background. Thus, the first half of the chapter is dedicated to giving a detailed account of perturbative quantum field theory, presented in a manner that, we hope, is palatable to the mathematician's taste. In particular, we discuss basic tools, such as the effective action and the perturbative expansion in Feynman graphs, as well as the regularization procedures used to evaluate the corresponding Feynman integrals. In particular, we concentrate on the procedure known as "dimensional regularization", both because of its being the one most commonly used in the actual calculations of particle physics, and because of the fact that it admits a very nice and conceptually simple interpretation in terms of noncommutative geometry, as we will come to see towards the end of the chapter. In this first half of Chapter 1 we give a new perspective on perturbative quantum field theory, which gives a clear mathematical interpretation to the *renormalization* procedure used by physicists to extract finite values from the divergent expressions obtained from the evaluations of the integrals associated to Feynman diagrams. This viewpoint is based on the Connes–Kreimer theory and then on more recent results by the authors.

Throughout this discussion, we always assume that we work with the procedure known in physics as "dimensional regularization and minimal substraction". The basic result of the Connes–Kreimer theory is then to show that the renormalization procedure corresponds exactly to the Birkhoff factorization of a loop $\gamma(z) \in G$ associated to the unrenormalized theory evaluated in complex dimension $D - z$, where D is the dimension of space-time and $z \neq 0$ is the complex parameter used in dimensional regularization. The group G is defined through its Hopf algebra of coordinates, which is the Hopf algebra of Feynman graphs of the theory. The Birkhoff factorization of the loop gives a canonical way of removing the singularity at $z = 0$ and obtaining the required finite result for the physical observables. This gives renormalization a clear and well defined conceptual meaning.

The Birkhoff factorization of loops is a central tool in the construction of solutions to the "Riemann-Hilbert problem", which consists of finding a differential equation with prescribed monodromy. With time, out of this original problem a whole area of mathematics developed, under the name of "Riemann–Hilbert correspondence". Broadly speaking, this denotes a way of encoding objects of differential geometric nature, such as differential systems with specified types of singularities, in terms of group representations. In its most general form, the Riemann–Hilbert correspondence is formulated as an equivalence of categories between the two sides. It relies on the "Tannakian formalism" to reconstruct the group from its category of representations. We give a general overview of all these notions, including the formalism of Tannakian categories and its application to differential systems and differential Galois theory.

The main new result of the first part of Chapter 1 is the explicit identification of the Riemann–Hilbert correspondence secretly present in perturbative renormalization. At the geometric level, the relevant category is that of *equisingular flat vector bundles.* These are vector bundles over a base space B which is a principal $\mathbb{G}_m(\mathbb{C}) = \mathbb{C}^*$-bundle

$$\mathbb{G}_m(\mathbb{C}) \to B \xrightarrow{\pi} \Delta$$

over an infinitesimal disk Δ. From the physical point of view, the complex number $z \neq 0$ in the base space Δ is the parameter of dimensional regularization, while the parameter in the fiber is of the form $\hbar\mu^z$, where $\hbar$ is the Planck constant and μ is a unit of mass. These vector bundles are endowed with a flat connection in the complement of the fiber over $0 \in \Delta$. The fiberwise action of $\mathbb{G}_m(\mathbb{C}) = \mathbb{C}^*$ is given by $\hbar\frac{\partial}{\partial\hbar}$. The equisingularity of the flat connection is a mathematical translation of the independence (in the minimal subtraction scheme) of the counterterms on the unit of mass μ. It means that the singularity of the connection, restricted to a section $z \in \Delta \mapsto \sigma(z) \in B$ of the bundle B, only depends upon the value $\sigma(0)$ of the section. We show that the category of equisingular flat vector bundles is a Tannakian category and we identify explicitly the corresponding group (more precisely, affine group scheme) that encodes, through its category of finite dimensional linear representations, the Riemann-Hilbert correspondence underlying perturbative renormalization. This is a very specific proalgebraic group of the form $\mathbb{U}^* = \mathbb{U} \rtimes \mathbb{G}_m$, whose unipotent part $\mathbb{U}$ is associated to the free graded Lie algebra

$$\mathcal{F}(1,2,3,\cdots)_{\bullet}$$

with one generator in each degree. We show that this group acts as a universal symmetry group of all renormalizable theories and has the properties of the "Cosmic Galois group" conjectured by Cartier. In many ways this group should be considered as the proper mathematical incarnation of the renormalization group whose role, as a group encoding the ambiguity inherent to the renormalization process in quantum field theory, is similar to that of the Galois group in number theory.

We conclude the first part of Chapter 1 with a very brief introduction to the theory of *motives* initiated by Grothendieck. We draw some parallels between the Tannakian formalism used in differential Galois theory and in particular in our formulation of perturbative renormalization and the same formalism in the context of motivic Galois groups. In particular we signal the fact that the group $\mathbb{U}^*$ also appears (albeit via a non-canonical identification) as a motivic Galois group in the theory of mixed Tate motives. This "motivic nature" of the renormalization group remains to be fully understood.

While the discussion in the first part of Chapter 1 applies to arbitrary renormalizable theories, the second part of this chapter is concerned with the theory which, as of the writing of this book, represents the best of our current knowledge of particle physics: the *Standard Model*. This part is based on joint work of the authors with Ali Chamseddine.

Our main purpose in the second part of Chapter 1 is to show that the intricate Lagrangian of the Standard Model minimally coupled to gravity, where we incorporate the terms that account for recent findings in neutrino physics, can be completely derived from very simple mathematical data. The procedure involves a modification of the usual notion of space-time geometry using the formalism of noncommutative geometry.

Again we do not assume that the reader has any familiarity with particle physics, so we begin this second part of Chapter 1 by reviewing the fundamental facts about the physics of the standard model and its coupling with gravity, in a formulation which is as close as possible to that of the physics literature. A main point that it is important to stress here is the fact that the standard model, in all its complexity, was built over the years as a result of a continuing dialogue between theory and experiment. The result is striking in its depth and complexity: even just the typesetting of the Lagrangian is in itself a time-consuming task.

After this introductory part, we proceed to give a brief description of the main tools of noncommutative geometry that will be relevant to our approach. They include cyclic and Hochschild cohomologies and the basic paradigm of *spectral triples* $(\mathcal{A},\mathcal{H},D)$. An important new feature of such geometries, which is absent in the commutative case, is the existence of inner fluctuations of the metric. At the level of symmetries, these correspond to the subgroup of inner automorphisms, a normal subgroup of the group of automorphisms which is non-trivial precisely in the noncommutative case.

We then begin the discussion of our model. This can be thought of as a form of unification, based on the symplectic unitary group in Hilbert space, rather than on finite dimensional Lie groups. The internal symmetries are unified with the gravitational ones. They all arise as automorphisms of the noncommutative algebra

of coordinates on a product of an ordinary Riemannian spin manifold M by a finite noncommutative space F. One striking feature that emerges from the computations is the fact that, while the metric dimension of F is zero, its K-theoretic dimension (in real K-theory) is equal to 6 modulo 8.
A long detailed computation then shows how the Lagrangian of the Standard Model minimally coupled with gravity is obtained naturally (in Euclidean form) from spectral invariants of the inner fluctuations of the product metric on $M \times F$.
This model provides specific values of some of the parameters of the Standard Model at unification scale, and one obtains physical predictions by running them down to ordinary scales through the renormalization group, using the Wilsonian approach. In particular, we find that the arbitrary parameters of the Standard Model, as well as those of gravity, acquire a clear geometric meaning in this model, in terms of moduli spaces of Dirac operators on the noncommutative geometry and of the asymptotic expansion of the corresponding *spectral action functional*. Among the physical predictions are relations between some of the parameters of the Standard Model, such as the merging of the coupling constants and a relation between the fermion and boson masses at unification.
Finally, in the last section of Chapter 1, we come to another application of noncommutative geometry to quantum field theory, which brings us back to the initial discussion of perturbative renormalization and dimensional regularization. We construct natural noncommutative spaces X_z of dimension a complex number z, where the dimension here is meant in the sense of the *dimension spectrum* of spectral triples. In this way, we find a concrete geometric meaning for the dimensional regularization procedure.
We show that the algebraic rules due to 't Hooft–Veltman and Breitenlohner–Maison on how to handle chiral anomalies using the dimensional regularization procedure are obtained, as far as one loop fermionic graphs are concerned, using the inner fluctuations of the metric in the product by the spaces X_z. This fits with the similar procedure used to produce the Standard Model Lagrangian from a product of an ordinary geometry by the finite geometry F and establishes a relation between chiral anomalies, computed using dimensional regularization, and the local index formula in NCG.
Towards the end of Chapter 1, one is also offered a first glance at the problem posed by a functional integral formulation of quantum gravity. We return only at the very end of the book to the problem of constructing a meaningful theory of quantum gravity, building on the experience we gain along the way through the analysis of quantum statistical mechanical systems arising from number theory, in relation to the statistics of primes and the Riemann zeta function. These topics form the second part of the book, to which we now turn.

The theme of Chapter 2 is the Riemann zeta function and its zeros. Our main purpose in this part of the book is to describe a spectral realization of the zeros as an absorption spectrum and to give an interpretation as a trace formula of the Riemann–Weil explicit formula relating the statistics of primes to the zeros of zeta. The role of noncommutative geometry in this chapter is twofold.
In the first place, the space on which the trace formula takes place is a noncommutative space. It is obtained as the quotient of the adèles $\mathbb{A}_{\mathbb{Q}}$ by the action of non-zero rational numbers by multiplication. Even though the resulting space $X = \mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$ is well defined set-theoretically, it should be thought of as a noncommutative space,

because the ergodicity of the action of $\mathbb{Q}^*$ on $\mathbb{A}_{\mathbb{Q}}$ prevents one from constructing measurable functions on the quotient X, as we show in Chapter 3. In particular, the construction of function spaces on X is done by homological methods using coinvariants. This will only acquire a full conceptual meaning in Chapter 4, using cyclic cohomology and the natural noncommutative algebra of coordinates on X.
The space X can be approximated by simpler spaces X_S obtained by restriction to finite sets S of places of $\mathbb{Q}$. We use this simplified setup to obtain the relation with the Riemann–Weil explicit formula. The main point is that, even though the space X_S is in essence a *product* of terms corresponding to the various places, the trace of the action of the group of idèle classes becomes a *sum* of such contributions. It is in the proof of this key additivity property that we use another tool of noncommutative geometry: the quantized calculus.
In the simplest instance, the interpretation of the Riemann–Weil explicit formula as a trace formula gives an interpretation as symplectic volume in phase space for the main term of the Riemann counting function for the asymptotic expansion of the number of non-trivial zeros of zeta of imaginary part less than E. We show that a full quantum mechanical computation then gives the complete formula.
We end Chapter 2 by showing how this general picture and methods extend to the zeta functions of arithmetic varieties, leading to a Lefschetz formula for the local L-factors associated by Serre to the Archimedean places of a number field.

The origin of the relation described above between the Riemann zeta function and noncommutative geometry can be traced to the work of Bost–Connes. This consists of the construction, using Hecke algebras, of a quantum statistical mechanical system whose partition function is the Riemann zeta function and which exhibits a surprising relation with the class field theory of the field $\mathbb{Q}$. Namely, the system admits as a natural symmetry group the group of idèle classes of $\mathbb{Q}$ modulo the connected component of the identity. This symmetry of the system is spontaneously broken at the critical temperature given by the pole of the partition function. Below this temperature, the various phases of the system are parameterized by embeddings $\mathbb{Q}^{cycl} \to \mathbb{C}$ of the cyclotomic extension $\mathbb{Q}^{cycl}$ of $\mathbb{Q}$. These different phases are described in terms of extremal KMS_β states, where $\beta = \frac{1}{T}$ is the inverse temperature. Moreover, another important aspect of this construction is the existence of a natural algebra of "rational observables" of this quantum statistical mechanical system. This allows one to define in a conceptual manner an action of the Galois group $\mathrm{Gal}(\mathbb{Q}^{cycl}/\mathbb{Q})$ on the phases of the system at zero temperature, merely by acting on the values of the states on the rational observables, values which turn out to provide a set of generators for $\mathbb{Q}^{cycl}$, the maximal abelian extension of $\mathbb{Q}$.
Our main purpose in Chapter 3 is to present extensions of this relation between number theory and quantum statistical mechanics to more involved examples than the case of rational numbers. In particular we focus on two cases. The first corresponds to replacing the role of the group GL_1 in the Bost–Connes (BC) system with GL_2. This yields an interesting non-abelian case, which is related to the Galois theory of the field of modular functions. The second is a closely related case of abelian class field theory, where the field $\mathbb{Q}$ is replaced by an imaginary quadratic extension. The results concerning these two quantum statistical mechanical systems are based, respectively, on work of the authors and on a collaboration of the authors with Niranjan Ramachandran.

We approach these topics by first providing a reinterpretation of the BC system in terms of geometric objects. These are the $\mathbb{Q}$-lattices, i.e. pairs (Λ, ϕ) of a lattice $\Lambda \subset \mathbb{R}^n$ (a cocompact free abelian subgroup of $\mathbb{R}^n$ of rank n) together with a homomorphism of abelian groups

$$\phi : \mathbb{Q}^n/\mathbb{Z}^n \longrightarrow \mathbb{Q}\Lambda/\Lambda.$$

Two $\mathbb{Q}$-lattices are commensurable if and only if

$$\mathbb{Q}\Lambda_1 = \mathbb{Q}\Lambda_2 \quad \text{and} \quad \phi_1 = \phi_2 \mod \Lambda_1 + \Lambda_2.$$

Let $\mathcal{L}_n$ denote the set of commensurability classes of n-dimensional $\mathbb{Q}$-lattices. Even in the simplest one-dimensional case ($n = 1$) the space $\mathcal{L}_n$ is a noncommutative space. In fact in the one-dimensional case it is closely related to the adèle class space $X = \mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$ discussed in Chapter 2.
We first construct a canonical isomorphism of the algebra of the BC system with the algebra of noncommutative coordinates on the quotient of $\mathcal{L}_1$ by the scaling action of $\mathbb{R}^*_+$. Following Weil's analogy between trigonometric and elliptic functions, we then show that the trigonometric analogue of the Eisenstein series generate, together with the commensurability with division points, the arithmetic subalgebra of "rational observables" of the BC system. This opens the way to the higher dimensional case and much of Chapter 3 is devoted to the extension of these results to the two-dimensional case.
The system for the GL_2 case is more involved, both at the quantum statistical level, where there are two phase transitions and no equilibrium state above a certain temperature, and at the number theoretic level, where the cyclotomic field $\mathbb{Q}^{cycl}$ is replaced by the modular field.
We end Chapter 3 with the description of our joint results with Ramachandran on the extension of the BC system to imaginary quadratic fields. This is based on replacing the notion of $\mathbb{Q}$-lattices with an analogous notion of 1-dimensional $\mathbb{K}$-lattices, with $\mathbb{K}$ the imaginary quadratic extension of $\mathbb{Q}$. The relation between commensurability of 1-dimensional $\mathbb{K}$-lattices and of the underlying 2-dimensional $\mathbb{Q}$-lattices gives the relation between the quantum statistical mechanical system for imaginary quadratic fields and the GL_2-system. This yields the relation between the quantum statistical mechanics of $\mathbb{K}$-lattices and the explicit class field theory of imaginary quadratic fields.

Underlying our presentation of the main topics of Chapter 3 there is a unifying theme. Namely, the three different cases of quantum statistical mechanical systems that we present in detail all fit into a similar general picture, where an ordinary moduli space is recovered as the set of classical points (zero temperature states) of a noncommutative space with a natural time evolution. In the setting of Chapter 3 the classical spaces are Shimura varieties, which can be thought of as moduli spaces of motives. This general picture will provide a motivating analogy for our discussion of the quantum gravity problem at the end of the book.

The spectral realization of zeros of zeta and L-functions described in Chapter 2 is based on the action of the idèle class group on the noncommutative space X of adèles classes. Nevertheless, the construction, as we describe it in Chapter 2, makes little use of the formalism of noncommutative geometry and no direct use of the crossed product algebra $\mathcal{A}$ describing the quotient of adèles by the multiplicative group $\mathbb{Q}^*$.

In Chapter 4, the last chapter of the book, we return to this theme. Our main purpose is to show that the spectral realization described in Chapter 2 acquires cohomological meaning, provided that one reinterprets the construction in terms of the crossed product algebra $\mathcal{A}$ and cyclic cohomology. This chapter is based on our joint work with Caterina Consani.

We begin the chapter by explaining how to reinterpret the entire construction of Chapter 2 in "motivic" terms using

- An extension of the notion of Artin motives to suitable projective limits, which we call *endomotives*.
- The category of cyclic modules as a linearization of the category of noncommutative algebras and correspondences.
- An analogue of the action of the Frobenius on ℓ-adic cohomology, based on a thermodynamical procedure, which we call *distillation*.

The construction of an appropriate "motivic cohomology" with a "Frobenius" action of $\mathbb{R}_+^*$ for endomotives is obtained through a very general procedure. It consists of three basic steps, starting from the data of a noncommutative algebra $\mathcal{A}$ and a state φ. One considers the time evolution $\sigma_t \in \operatorname{Aut}\mathcal{A}$, for $t \in \mathbb{R}$, naturally associated to the state φ.

The first step is what we refer to as *cooling*. One considers the space $\mathcal{E}_\beta$ of extremal KMS_β states, for β greater than critical. Assuming these states are of type I, one obtains a morphism

$$\pi : \mathcal{A} \rtimes_\sigma \mathbb{R} \to \mathcal{S}(\mathcal{E}_\beta \times \mathbb{R}_+^*) \otimes \mathcal{L}^1,$$

where $\mathcal{A}$ is a dense subalgebra of a C^*-algebra $\bar{\mathcal{A}}$, and where $\mathcal{L}^1$ denotes the ideal of trace class operators. In fact, one considers this morphism restricted to the kernel of the canonical trace τ on $\bar{\mathcal{A}} \rtimes_\sigma \mathbb{R}$.

The second step is *distillation*, by which we mean the following. One constructs a cyclic module $D(\mathcal{A}, \varphi)$ which consists of the cokernel of the cyclic morphism given by the composition of π with the trace $\mathrm{Tr} : \mathcal{L}^1 \to \mathbb{C}$.

The third step is then the *dual action*. Namely, one looks at the spectrum of the canonical action of $\mathbb{R}_+^*$ on the cyclic homology

$$HC_0(D(\mathcal{A}, \varphi)).$$

This procedure is quite general and it applies to a large class of data $(\mathcal{A}, \varphi)$, producing spectral realizations of zeros of L-functions. It gives a representation of the multiplicative group $\mathbb{R}_+^*$ which combines with the natural representation of the Galois group G when applied to the noncommutative space (analytic endomotive) associated to an (algebraic) endomotive.

In the particular case of the endomotive associated to the BC system, the resulting representation of $G \times \mathbb{R}_+^*$ gives the spectral realization of the zeros of the Riemann zeta function and of the Artin L-functions for abelian characters of G. One sees in this example that this construction plays a role analogous to the action of the Weil group on the ℓ-adic cohomology. It gives a functor from the category of endomotives to the category of representations of the group $G \times \mathbb{R}_+^*$. Here we think of the action of $\mathbb{R}_+^*$ as a "Frobenius in characteristic zero", hence of $G \times \mathbb{R}_+^*$ as the corresponding Weil group.

We also show that the "dualization" step, i.e. the transition from $\mathcal{A}$ to $\mathcal{A} \rtimes_\sigma \mathbb{R}$, is a very good analog in the case of number fields of what happens for a function field K in passing to the extension $K \otimes_{\mathbb{F}_q} \bar{\mathbb{F}}_q$. In fact, in the case of positive characteristic,

the unramified extensions $K \otimes_{\mathbb{F}_q} \mathbb{F}_{q^n}$, combined with the notion of places, yield the points $C(\bar{\mathbb{F}}_q)$ over $\bar{\mathbb{F}}_q$ of the underlying curve. This has a good parallel in the theory of factors and this analogy plays an important role in developing a setting in noncommutative geometry that parallels the algebro-geometric framework that Weil used in his proof of RH for function fields.
We end the number theoretic part of the book by a dictionary between Weil's proof and the framework of noncommutative geometry, leaving open the problem of completing the translation and understanding the noncommutative geometry of the "arithmetic site".

We end the book by coming back to the construction of a theory of quantum gravity. Our approach here starts by developing an analogy between the electroweak phase transition in the Standard Model and the phase transitions in the quantum statistical mechanical systems described in Chapter 3. Through this analogy a consistent picture emerges which makes it possible to define a natural candidate for the algebra of observables of quantum gravity and to conjecture an extension of the electroweak phase transition to the full gravitational sector, in which the geometry of space-time emerges through a symmetry breaking mechanism and a cooling process. As a witness to the unity of the book, it is the construction of the correct category of correspondences which, as in Grothendieck's theory of motives, remains the main challenge for further progress on both QG and RH.

Acknowledgment. We wish to thank our friends and collaborators Ali Chamseddine, Katia Consani, Michel Dubois-Violette, Dirk Kreimer, Yuri Manin, Henri Moscovici, and Niranjan Ramachandran, whose ideas and contributions are reflected in much of the content of this book. We thank Jorge Plazas for accepting to be the official "test reader" of the book and for the many useful comments he provided. We also thank all the people who offered to send us comments and suggestions, among them Pierre Cartier, Masoud Khalkhali, Peter May, Laura Reina, Abhijnan Rej, Thomas Schücker, Walter van Suijlekom, Joseph Varilly and Don Zagier. Various institutions provided hospitality and support at various stages during the preparation of the book: we thank the IHES and the MPI for making several mutual visits possible, as well as the Kavli Institute, the Newton Institute, and Vanderbilt University for some very productive extended stays. We thank Arthur Greenspoon and Patrick Ion for a very careful proofreading of the manuscript and the editors at both the AMS and HBA for their infinite patience. This work was partially supported by NSF grants DMS-0652164 and DMS-0651925.

CHAPTER 1

Quantum fields, noncommutative spaces, and motives

1. Introduction

The main goal of Chapter one is to unveil the mathematical conceptual meaning of some of the sophisticated computations performed by physicists in the domain of particle physics. It is divided into two parts dealing, respectively, with

(1) Renormalization
(2) The Standard Model

We try to keep close contact with the way the computations are actually performed by physicists and to bridge the gap between, on the one hand, the lessons physicists learned from their constant dialogue with experimental results and, on the other hand, with the elaborate mathematical concepts involved in allowing one to understand the meaning of these computations (if any). The bare data we start with are, respectively,

- Computations of cross sections and scattering amplitudes from the perturbative expansion of the Feynman integral using renormalized values of Feynman graphs in the dimensional regularization and minimal subtraction scheme.
- The detailed expression of the full Standard Model Lagrangian with neutrino mixing, the see-saw mechanism and coupling to gravity.

We start the chapter with a presentation of quantum field theory (QFT) that ought to be understandable to mathematicians. In particular we recall in §2.1 the Lagrangian and Hamiltonian formalisms of classical mechanics and explain in §2.2 how the Lagrangian formulation of quantum field theory leads to Feynman's path integral. The path integral prescription is ambiguous (even ignoring all the divergence problems) and the removal of the ambiguity by Feynman's $i\epsilon$ trick can only be properly understood after an excursion into the Hamiltonian formulation and canonical quantization. We do this in §2.3 where we base the discussion on the three main physical properties of a quantum field theory, which are

- Causality
- Positivity of energy
- Unitarity

We discuss the simplest example of QFT in §2.4: it is the free bosonic field on the space-time

$$X = \mathbb{R} \times S^1$$

with the Lorentz metric. This gives a good occasion to describe this basic example of algebraic quantum field theory, and to explain what are the vacuum states and the temperature states which fulfill the KMS_β condition relative to the Heisenberg time evolution. Even though algebraic quantum field theory is an interesting formalism

involving deep mathematical structures such as von Neumann algebras it falls short of what is essential to our purpose: concrete physics computations. It is, however, essential in clarifying the conceptual meaning of the boundary conditions on Green's functions by clearly separating the kinematical relations from the construction of the vacuum states. We turn to the Green's functions in §2.5 and give their formal perturbative expansion in terms of free fields as the Gell-Mann–Low formula. . We then show in §2.6 how Wick rotation allows one to encompass Feynman's $i\epsilon$ prescription for how to go around the pole of the propagator in the analytic continuation to imaginary time. We give all the details on that point since it removes the first ambiguity in the perturbative computation of the functional integral and shows from the start the merit of the Euclidean formulation.
The Feynman graphs are dealt with in §3. We start with a detailed account of a concrete example in §3.1 and show how the various pairings coming from the integration by parts under a Gaussian are labeled by graphs and yield integrals. The simplest graphs, such as self-energy graphs, give rise to integrals which diverge when an ultraviolet cutoff is removed so that one is confronted with the problem of renormalization. The physics origin of the problem was already understood by Green in 1830 and we explain the computation of the self-energy in hydrodynamics in §3.2 as a first example of mass renormalization. We then use the analogy between hydrodynamics and electromagnetism to explain how the crucial distinction between the bare parameters and the observed ones makes it possible to eliminate the divergence of the simplest self-energy graph by adding counterterms to the Lagrangian. We give a precise mathematical definition of Feynman graphs and of the rules which associate a formal integral to a graph in §3.3. We then describe the standard procedures that allow one to simplify the combinatorics of the Feynman graphs. First, by taking the logarithm of the partition function with a source term one reduces to connected graphs (§3.4). Then after applying the Legendre transform one obtains the effective action. Both the action's role as the basic unknown of QFT and its expansion in terms of one-particle irreducible (1PI) graphs is explained in section §3.5. With this tool at hand we come to a precise definition of the physical parameters, such as the mass, and observables, such as the scattering matrix, in terms of the effective action in §3.6. Finally we describe the physical ideas of mass, field strength and coupling constant renormalization in §3.7.
In §4 we recall the basic dimensional regularization and minimal subtraction procedures (DimReg+MS). We begin with the very simple example of the self-energy graph for the scalar ϕ^3 theory. We show explicitly, in this example, how to implement the dimensional regularization of the divergent integral using Schwinger parameters and the formal rules for Gaussian integration in a complexified dimension $D - z$. We then discuss the existence of an analytic continuation of the Feynman integrals to a meromorphic function on the complex plane. We prove in Theorem 1.9 that the dimensionally regularized unrenormalized values $U^z(\Gamma(p_1, \ldots, p_N))$ have the property that their Taylor coefficients at $p = 0$ admit a meromorphic continuation to the whole complex plane $z \in \mathbb{C}$. In §5.1 we show a simple example of a *subdivergence* for a 1PI (one-particle irreducible) graph of the massless ϕ^3 theory in dimension 6. This example shows the need, in addition to the regularization scheme (here DimReg+MS), for a renormalization procedure that accounts for the combinatorics of nested subdivergences.

In §5, we introduce the Bogoliubov–Parasiuk–Hepp–Zimmermann renormalization (BPHZ) procedure. This takes care of eliminating the divergences step by step in the perturbative expansion, by repeatedly adjusting the bare parameters in the Lagrangian by suitable counterterms that cancel the divergences. The BPHZ procedure also takes care of the problem of non-local counterterms associated to subdivergences. We show this in detail in an explicit example in §5.1. We also explain the role of the external structure of Feynman graphs. The counting of the degree of divergence is described in §5.2. The BPHZ preparation of Feynman graphs and the extraction of the renormalized value and the inductive definition of the counterterms are discussed in §5.3.

In §6.1 we give some mathematical background on commutative Hopf algebras and affine group schemes, while in § 6.2 we introduce the Connes–Kreimer Hopf algebra of Feynman graphs, first only in its discrete combinatorial version. Then in §6.3 we refine the construction, by taking also into account the external structure. Theorem 1.39 gives the recursive formula of Connes–Kreimer for the Birkhoff factorization in a graded connected Hopf algebra, which gives a clear conceptual interpretation to the BPHZ procedure, when applied to the Hopf algebra of Feynman graphs.

In §6.5 we recall another result of the Connes–Kreimer theory, relating the affine group scheme of the Hopf algebra of Feynman graphs, called the group of diffeographisms of the physical theory, to formal diffeomorphisms of the coupling constants of the theory. We explain in §6.6 the dependence of the $U^z(\Gamma(p_1, \ldots, p_N))$ on a mass parameter μ and how to recover in the Connes–Kreimer theory the notion of renormalization group lifted to the level of the group of diffeographisms, with the β-function given by an element in the corresponding Lie algebra.

This singles out the data of perturbative renormalization as describing a certain class of loops in the affine group scheme of diffeographisms, satisfying some conditions on the dependence on the mass parameter μ, with the renormalization procedure consisting of their Birkhoff factorization.

We give in §7 a reinterpretation of these data in terms of a Riemann–Hilbert correspondence. We begin in §7.1 with the expression of the counterterms as a "time-ordered exponential" (iterated integral). We then introduce in §7.2 the notion of flat equisingular connection, which reformulates geometrically the conditions satisfied by the class of loops corresponding to the data of perturbative renormalization. The corresponding equivariant principal bundles are described in §7.3.

In §7.4 we recall some general facts about Tannakian categories, the Tannakian formalism, and representations of affine group schemes. In §7.5 we show how this formalism is variously used in the context of differential Galois theory to classify categories of differential systems with prescribed singularities through the Riemann–Hilbert correspondence, a broad generalization of the original Riemann–Hilbert problem on the reconstruction of differential equations from their monodromy representations.

In §7.6 we apply this general strategy to the case of renormalization. We introduce a category of flat equisingular vector bundles, and we obtain in Theorem 1.100 an identification with the category $\mathrm{Rep}_{\mathbb{U}^*}$ of finite-dimensional linear representations of the affine group scheme $\mathbb{U}^* = \mathbb{U} \rtimes \mathbb{G}_m$ with Hopf algebra $\mathcal{H}_{\mathbb{U}} := U(\mathcal{F}(1, 2, 3, \cdots)_\bullet)^\vee$ where $\mathcal{F}(1, 2, 3, \cdots)_\bullet$ is the free graded Lie algebra with one generator e_{-n} in each degree $n > 0$.

In §8 we give a brief introduction to the theory of motives, which plays a role in many different ways throughout the book. In particular we mention the fact that the affine group scheme $\mathbb{U}^*$ also has an incarnation as a motivic Galois group for a category of mixed Tate motives. Of the general aspects of the theory of motives, we recall briefly the general interpretation as a universal cohomology theory, the construction of the category of pure motives, the role of algebraic cycles and equivalence relations, the relation to zeta functions (which plays an important role in Chapter 2), the Weil conjectures and the Grothendieck standard conjectures, the role of the Tannakian formalism and motivic Galois groups, the special case of Artin motives (which plays a role in Chapter 4), mixed motives and, in particular, the mixed Tate motives that seem to be deeply related to quantum field theory and their relation to mixed Hodge structures.

This completes the first part of Chapter 1. In the second part of the chapter we deal with the Standard Model of elementary particle physics and an approach to a simple mathematical understanding of its structure via noncommutative geometry. The second part of Chapter 1 follows closely our joint work with Chamseddine [**52**], which is based on the model introduced by Connes in [**73**], as well as on the previous work of Chamseddine and Connes on the spectral action [**45**], [**46**], [**47**].
Since we do not assume that the reader has much familiarity with particle physics, we begin §9 by giving a brief overview of the *Standard Model*. We introduce the various parameters, the particles and interactions, symmetries in §§9.1, 9.2, 9.3 and we reproduce in full in §9.4 the very complicated expression of the Lagrangian. We discuss in §9.5 other aspects, such as the ghost terms and gauge fixing, that become relevant at the quantum rather than the semi-classical level. In § 9.6 we distinguish between the *minimal* Standard Model, which has only left-handed massless neutrinos, and the extension that is required in order to account for the experimental evidence of neutrino mixing. We describe in §9.6.2 the corresponding modifications of the Lagrangian. In §§9.7 and 9.8 we describe the Lagrangian that gives the Standard Model minimally coupled to gravity, where the gravity part can be considered as an effective field theory by including higher-derivative terms.
The problem of realizing the symmetries of particle interactions as diffeomorphisms (pure gravity) on a suitable space suggests the idea that noncommutative geometry, where inner symmetries are naturally present, should provide the correct framework. We discuss this in §9.9.
In §10 we then recall the main notions of (metric) noncommutative geometry developed by Connes, based on the structure of *spectral triple* that generalizes Riemannian geometry to the noncommutative setting. We introduce spectral geometry in §10.1, we recall the definition and basic properties of spectral triples in §10.2, including the *real part* defined in §10.3. We recall some well known facts on Hochschild and cyclic cohomology in §10.4. The local index formula of Connes–Moscovici is recalled briefly in §10.5. The Yang–Mills and Chern–Simons actions are described in terms of Hochschild cohomology in §10.6 and §10.7 following work of Chamseddine–Connes.
The important notion of inner fluctuations of the metric associated to self-Morita equivalences of a noncommutative space is discussed in §10.8.
In §11 we introduce the *spectral action principle* of Chamseddine–Connes. This plays a crucial role in recovering the Standard Model Lagrangian from noncommutative geometry and is one of the main tools in metric noncommutative geometry. A careful

discussion of the terms arising in the asymptotic expansion of the spectral action functional is given in §§11.1, 11.2, 11.3 using Seeley–DeWitt coefficients and Gilkey's theorem.

In §11.4 we recall a result of [**46**] that illustrates how to recover the Einstein–Yang–Mills action from the spectral action on the very simple noncommutative space given by the product of an ordinary 4-manifold by a noncommutative space described by the algebra $M_N(\mathbb{C})$ of $N \times N$ matrices.

We then analyze the terms that appear in the asymptotic expansion of the spectral action in §11.5 and their behavior under inner fluctuations of the metric. We also recall briefly in §11.6 the modification of the spectral action by a *dilaton field* introduced by Chamseddine and Connes in [**47**].

In §12 we begin the discussion of the noncommutative geometry of the Standard Model, following [**52**]. We introduce in §13 a finite noncommutative geometry F, derived from the basic input of the model, which is the finite-dimensional associative algebra $\mathbb{C} \oplus \mathbb{H} \oplus \mathbb{H} \oplus M_3(\mathbb{C})$, with $\mathbb{H}$ the real division algebra of quaternions. We explain how to construct canonically an *odd spin* representation $\mathcal{H}_F$ of this algebra and, upon imposing a very natural condition on possible Dirac operators for this geometry, we identify in §13.1 a maximal subalgebra compatible with the existence of Dirac operators with the required properties. The subalgebra is of the form

$$\mathcal{A}_F = \mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C}) \subset \mathbb{C} \oplus \mathbb{H} \oplus \mathbb{H} \oplus M_3(\mathbb{C}).$$

In §13.2 we identify the bimodule $\mathcal{H}_F$ with the fermions of the Standard Model (after fixing the number of generations $N = 3$) and we show in §13.3 that this identification is dictated by the fact that it reproduces the correct values of the hypercharges.

In §§13.4 and 13.5 we give a complete classification for the possible Dirac operators for this finite geometry and we describe their moduli space, which gives a geometric interpretation for the Yukawa parameters of the Standard Model. The intersection pairing of the finite geometry is analyzed in §13.6, using the fact that the metric and KO-dimensions are not the same, the first being zero and the second being equal to 6 modulo 8.

We then consider in §14 the product $M \times F$ of an ordinary compact spin 4-manifold with the finite noncommutative geometry introduced previously, described as a cup product of spectral triples. We identify the real part of the product geometry in §14.1.

The bosons of the Standard Model, including the Higgs field, are obtained as inner fluctuations of the metric on the product geometry in §15, with the discrete part giving the Higgs analyzed in §15.2 and the gauge bosons in §15.4.

The main computation that shows how to recover the Standard Model Lagrangian, including mixing and Majorana mass terms for neutrinos, minimally coupled to gravity, is carried out in §16 by breaking down the Lagrangian in several steps and relating the resulting terms to the terms in the asymptotic expansion of the spectral action functional

$$\mathrm{Tr}(f(D/\Lambda)) + \frac{1}{2}\langle J\psi, D\psi\rangle,$$

with the additional fermionic term $\langle J\psi, D\psi\rangle$. In particular, we explain in §§16.2 and 16.3 how the fermionic term gives rise to a Pfaffian which takes care of the "fermion doubling problem" of [**210**] by taking the square root of a determinant. Among the physical consequences of deriving the Lagrangian from the spectral action, and making the "big desert" hypothesis, we find in § 17.2 the merging of the coupling constants at unification, in the form of the relation $g_2 = g_3 = \sqrt{5/3}\, g_1$ typical of the grand unified theories. We also find in §§17.4 and 17.6 a simple quadratic relation between the masses of quarks and leptons and the W-mass at unification, compatible with the known physics at ordinary energies. In §17.5 and §17.10 we show that this model also provides a *see-saw mechanism* that accounts for the observed smallness of neutrino masses, and a prediction of a heavy Higgs mass at around 168 GeV. The gravitational terms are discussed in §17.11 and the geometric interpretation of the free parameters of the Standard Model is summarized in §17.12.

In §18 we outline a possible functional integral formulation in the context of spectral geometry. We briefly explain a more conceptual path to the algebra $\mathcal{A}_F$ and its representation given in [**49**], [**50**] based on the classification of irreducible finite geometries of KO-dimension 6 modulo 8. This quickly leads to the algebra $M_2(\mathbb{H}) \oplus M_4(\mathbb{C})$ with a non-trivial grading on $M_2(\mathbb{H})$ and to its natural representation, playing the role of the above odd-spin representation. The same mechanism, coming from the order one condition, then reduces to the subalgebra

$$\mathcal{A}_F = \mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C}) \subset M_2(\mathbb{H}) \oplus M_4(\mathbb{C}).$$

We also show that the use of the larger algebra restores both the orientability and the Poincaré duality.

In the remaining part of Chapter 1, in §19 we return to the dimensional regularization procedure described in the first part of the chapter in the context of the Connes–Kreimer theory. Here we give the construction, using spectral triples, of a noncommutative space X_z whose dimension (in the sense of the dimension spectrum of a spectral triple) consists of a complex number z, and we reinterpret geometrically the DimReg procedure, as far as one-loop fermionic graphs are concerned, as taking the cup product of the spectral triple associated to an ordinary manifold of integer dimension D with this noncommutative space. In Chapter 2 we return to discuss this construction and we give an arithmetic model of a noncommutative space X_z of dimension z. We show in §19.1 and following that the construction of X_z is compatible with the Breitenlohner–Maison prescription for treating γ_5 in the context of DimReg. We continue with the discussion of anomalies, by introducing chiral gauge transformations in §19.3, and discussing the finiteness of the anomalous graphs in §19.4. We treat explicitly the simplest cases of anomalous graphs in §19.5 and we relate anomalous graphs in dimension 2 and the local index cocycle in §19.6.

2. Basics of perturbative QFT

Quantum field theory is the most accurate source of predictions about the world of elementary particles. At the theoretical level, it is full of subtleties and ingenious procedures that extract finite and experimentally testable values from formal series of divergent integrals. The development of this theory, which achieved the unification of two fundamental revolutions of early twentieth century physics, special relativity and quantum mechanics, traces its origins to two crucial developments that took place in the late 1920s. The quantization of the electromagnetic field by Born, Heisenberg, and Jordan in 1925 explained Einstein's fluctuation formula for the electromagnetic radiation in a cavity. It was soon followed by Dirac's foundational paper [**119**] of 1927 which performed the first quantum electrodynamics calculation and explained Einstein's ansatz for the probabilities of spontaneous and induced transitions in an atom. The second crucial development is Dirac's relativistic wave

equation for the electron which gave the spin, the magnetic moment, the fine structure of the atomic spectra and predicted the positron. The advent of quantum field theory provided a formalism in which particles and fields are treated on an equal footing. This is crucial in order to describe phenomena involving the interaction of radiation and matter, as well as all the particle physics events involving collisions, where typically the number of particles is not preserved. In such events, in fact, particles are created and annihilated, via the conversion of energy into mass and vice versa. This is a typically relativistic phenomenon that cannot be accounted for simply by the use of quantum mechanics.
The problem of divergences in quantum field theory manifested itself very early, in computing the self-energy of the interaction of a charged point particle with its own electromagnetic field. It became clear in the early stages of development of the theory that the problem of divergences was a substantial difficulty that systematically plagued the calculations. Typically, the problem arises from the fact that, at very high energies, there is an increasingly large number of virtual states to sum over, so that the resulting series are divergent. The fact that, for a "renormalizable" theory, one can cure these infinities is a manifestation of the idea that the corresponding physical processes are insensitive to the virtual states that live above a certain energy scale.

2.1. Lagrangian and Hamiltonian formalisms.

We recall for convenience the relation between the Lagrangian and Hamiltonian formulations of classical mechanics.
Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_j}\right) - \frac{\partial L}{\partial q_j} = 0 \tag{1.1}$$

follow from the stationary value of the action integral

$$S = \int L(q,\dot q)\, dt\,. \tag{1.2}$$

The simplest example of a Lagrangian is of the form

$$L(q,\dot q) = \frac{1}{2} m \dot q^2 - V(q) \tag{1.3}$$

where $V(q)$ is a potential function.
The Lagrangian $L(q,\dot q)$ is a function on the tangent space TC of the configuration space C. The Hamiltonian is a function on the cotangent space T^*C obtained as the Legendre transform of L. Thus at a given point $q \in C$ one considers the map λ from the tangent space T_qC to the cotangent space T_q^*C defined by differentiating the restriction of L to T_qC. The differential $\partial_{\dot q} L$ is a linear form on T_qC and hence an element of T_q^*C, thus

$$\lambda(\dot q) = \partial_{\dot q} L(q,\dot q) \in T_q^*C\,. \tag{1.4}$$

The inverse correspondence λ^{-1} from the cotangent space T_q^*C to the tangent space T_qC is obtained by assigning to $p \in T_q^*C$ the values of $\dot q \in T_qC$ for which the following expression is stationary in $\dot q$,

$$\dot q \mapsto p\,\dot q - L(q,\dot q)\,. \tag{1.5}$$

There is no guarantee that λ is invertible. When it is one gets a function, the Hamiltonian, on T_q^*C defined by

$$H(p,q) = p\,\lambda^{-1}(p) - L(q,\lambda^{-1}(p))\,, \quad \forall p \in T_q^*C\,. \tag{1.6}$$

In other words the Hamiltonian is given by the value of (1.5) at the (unique by hypothesis) stationary point.

One can now apply the same Legendre transformation to the Hamiltonian, i.e. consider the map λ' from the cotangent space T_q^*C to the tangent space T_qC defined by differentiating the restriction of H to T_q^*C. In fact one has $\lambda' = \lambda^{-1}$. Indeed, when we differentiate (1.6) with respect to p the stationarity condition (1.5) shows that the only contribution is the coefficient $\lambda^{-1}(p)$ of p in the first term.

In other words the relation between the Lagrangian and the Hamiltonian is symmetric; each is the Legendre transform of the other. Notice also that the important expression is $L\,dt$, which can be written in the form

$$L(q,\dot q)\,dt = p\,dq - H\,dt \tag{1.7}$$

which already brings in the Lorentz signature.

For functions on the cotangent space T^*C the Poisson bracket is defined by

$$\{f,g\} = \left\langle \frac{\partial f}{\partial p}, \frac{\partial g}{\partial q} \right\rangle - \left\langle \frac{\partial f}{\partial q}, \frac{\partial g}{\partial p} \right\rangle \tag{1.8}$$

where the pairings take place as above between the tangent and cotangent spaces at the same point. The equations of motion (1.1) in Hamiltonian form just give the evolution of any function f on the cotangent space T^*C,

$$\frac{d}{dt} f = \{H, f\} \tag{1.9}$$

In particular H itself is a constant of the motion, which is the energy of the system. One has to take care of overall signs (which were unimportant insofar as the equations of motion were concerned) since stability of a system means that the energy is bounded below. For the simplest Lagrangians such as (1.3) the Hamiltonian is simply

$$H(p,q) = \frac{p^2}{2m} + V(q) \tag{1.10}$$

and the sign of the potential function is the *opposite* one of the one it had in the Lagrangian.

One also needs to pay attention to “units”. We use the speed of light as a conversion factor from time to length and thus set $c = 1$. Thus length and time are on the same scale T (i.e. they have the same physical dimensions). Similarly, masses, energies and momenta are on the same scale E. Both the Lagrangian L and the Hamiltonian have physical dimension E. The action (1.2) has physical dimension $E \cdot T$, which is quantized in units of $\hbar$. One could set $\hbar = 1$ but it is useful to keep track of powers of $\hbar$ in the perturbative expansions in QFT.

2.2. Lagrangian and the Feynman integral.

A quantum field theory in dimension D (one usually assumes $D = 4$, but we will see a significant example with $D = 6$) is specified by assigning a classical action functional

$$S(\phi) = \int \mathcal{L}(\phi)\, d^D x, \tag{1.11}$$

where ϕ is a classical field and the Lagrangian density is of the form

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\,\phi^2 - \mathcal{L}_{\rm int}(\phi). \tag{1.12}$$

Here we write $(\partial\phi)^2 = (\partial_0\phi)^2 - \sum_{\nu\neq 0}(\partial_\nu\phi)^2$, with Lorentzian signature, i.e.

$$(\partial\phi)^2 = g^{\mu\nu}\,\partial_\mu\phi\,\partial_\nu\phi$$

where

$$(g^{\mu\nu})_{\mu,\nu=1,\ldots,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.13}$$

The first two terms

$$\mathcal{L}_0(\phi) = \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\,\phi^2 \tag{1.14}$$

give the free field part of the Lagrangian, while the remaining term $\mathcal{L}_{\rm int}(\phi)$ is the interaction part, which is usually assumed to be a polynomial in the field ϕ.

In terms of physical units, since we observed that (1.11) has the physical dimensions $E\cdot T$, we see from (1.12) that the field ϕ is expressed in physical units of $T^{(D-2)/2}$. We say in this case simply that the field ϕ has physical dimension $D/2-1$. Similarly, one can deduce from (1.12) the physical dimensions of the coupling constants that appear in the term $\mathcal{L}_{\rm int}(\phi)$.

Definition 1.1. *A theory $\mathcal{T}$ is specified by assigning the data of a Lagrangian density* (1.12) *and a positive integer D, the dimension.*

A classical observable is a function $\mathcal{O}(\phi)$ of the classical fields. One can *define* the quantum expectation value of a classical observable by the expression

$$\langle\mathcal{O}\rangle = \mathcal{N}\int \mathcal{O}(\phi)\, e^{i\,\frac{S(\phi)}{\hbar}}\,\mathcal{D}[\phi], \tag{1.15}$$

where $\mathcal{N}$ is a normalization factor,

$$\mathcal{N}^{-1} = \int e^{i\,\frac{S(\phi)}{\hbar}}\,\mathcal{D}[\phi].$$

Notice that the (Feynman) integral here has only a formal meaning, since the "measure" $\mathcal{D}[\phi]$ is typically not well defined mathematically. However, this suffices in the case where the space of classical fields ϕ is a linear space, in order to define without difficulty the terms in the perturbative expansion, which make the renormalization problem manifest.

Notice also how the expression (1.15) formally resembles the averages that are used in computing observables and correlation functions in statistical mechanics, except for the important difference that here, in the transition from "classical field theory" to "quantum field theory", one replaces the classical notion of probabilities by complex-valued *probability amplitudes* and asserts that the probability amplitude of a classical field configuration ϕ is given by the formula of Dirac and Feynman

$$\exp\left(i\,\frac{S(\phi)}{\hbar}\right), \tag{1.16}$$

where $S(\phi)$ is the classical action (1.11) and $\hbar$ is the unit of action, so that $iS(\phi)/\hbar$ is a dimensionless quantity.

2.3. The Hamiltonian and canonical quantization.

In classical field theory the configuration space C is the space of fields $\phi(x^j)$ viewed as functions on "space" (one uses traditionally the Greek indices for space-time variables with $x^0 = t$ and Latin indices for the space variables). A space-time field $\phi(x^\mu)$ is viewed as a trajectory in C

$$t \mapsto q(t) = \phi(t, \bullet) \in C\,.$$

The classical action (1.2) is of the form

$$S(\phi) = \int L(q, \dot q)\, dt$$

where $L(q, \dot q)$ is of the form (1.3). The kinetic term is given by the functional

$$\frac{1}{2}\int \dot\phi(x)^2\, d^{D-1}x$$

on TC, where we use the linearity of C to identify it with its tangent space at a given point. The potential V is given by the function

$$V(\phi) = \int \left(\frac{1}{2}\,(\partial_j\phi)^2 + \frac{m^2}{2}\,\phi^2 + \mathcal{L}_{\rm int}(\phi)\right)\, d^{D-1}x$$

on C and the Hamiltonian is

$$H(\pi, \phi) = \int \left(\frac{1}{2}\,\pi^2 + \frac{1}{2}\,(\partial_j\phi)^2 + \frac{m^2}{2}\,\phi^2 + \mathcal{L}_{\rm int}(\phi)\right)\, d^{D-1}x\,. \tag{1.17}$$

Canonical quantization replaces $\phi(x)$ and $\pi(x)$ by noncommuting variables $\tilde\phi(x)$ and $\tilde\pi(x)$ which for a simple scalar bosonic field as above fulfill the canonical commutation relations

$$[\tilde\phi(x), \tilde\phi(y)] = 0\,, \quad [\tilde\pi(x), \tilde\pi(y)] = 0\,, \quad [\tilde\pi(x), \tilde\phi(y)] = -\,i\,\hbar\,\delta(x-y)\,. \tag{1.18}$$

One then defines the quantum field $\tilde\phi(t,x)$ at the formal level by the basic formula

$$\tilde\phi(t,x) = e^{i\,t\,H}\,\tilde\phi(x)\,e^{-i\,t\,H} \tag{1.19}$$

where H stands for the evaluation of the function $H(\pi,\phi)$ of (1.17) on the noncommuting variables $\tilde\phi(x)$ and $\tilde\pi(x)$. This is a highly ambiguous and ill-defined element of the algebra $\mathcal{A}$ generated by these variables. What matters is the one-parameter group

$$\sigma_t(T) = e^{i\,t\,H}\,T\,e^{-i\,t\,H} \tag{1.20}$$

giving the time evolution of observable quantities. Once one has succeeded in making sense of this group the main problem is to classify the vacuum states, i.e. the representations ρ of the algebra $\mathcal{A}$ in a Hilbert space $\mathcal{H}$ such that

- ρ is an irreducible representation
- There exists a *positive* unbounded operator H in $\mathcal{H}$ such that
$$\rho(\sigma_t(T)) = e^{itH}\,\rho(T)\,e^{-itH}\,, \quad \forall T \in \mathcal{A}$$
- There exists a unique eigenvector $|\,0\,\rangle \in \mathcal{H}$, such that $H|\,0\,\rangle = 0$ and the subspace $\rho(\mathcal{A})|\,0\,\rangle \subset \mathcal{H}$ is dense in $\mathcal{H}$

Such representations are characterized by the positive linear functional on $\mathcal{A}$:
$$\psi(T) = \langle\, 0|\,\rho(T)|\,0\,\rangle$$
whose main property is to be a pure state and to fulfill the following limiting case of the Kubo–Martin–Schwinger (KMS) condition (which will play a crucial role in Chapter 3 below):

- For bounded observables $a, b \in \mathcal{A}$ the correlation function
$$F(t) = \psi(a\,\sigma_t(b))$$
is the boundary value of a bounded holomorphic function $F(z)$ in the upper half-plane.

In practice it is very hard to make sense of the one-parameter group σ_t and then to classify the vacuum states. It is nevertheless quite important to keep in mind that there are two separate steps that need to be taken in order to obtain a quantum field theory:

- The definition of the observables and the dynamics.
- The construction and classification of vacuum states.

The three main physical properties of a quantum field theory are

- Causality
- Positivity of energy
- Unitarity

The first two are already incompatible in quantum mechanics [**130**].

Causality is a property of the dynamics; its best formulation says that (for bosonic fields) the quantum fields commute at space-like separation,

$$[\tilde{\phi}(x),\, \tilde{\phi}(y)] = 0\,, \quad \forall (x, y)\,, \quad (x - y)^2 < 0, \tag{1.21}$$

where two points x, y of space-time are space-like separated iff $(x - y)^2 < 0$ in the metric (1.13).

Unitarity is equivalent to the property of the vacuum state that

$$\psi(a^*\,a) \geq 0\,, \quad \forall a \in \mathcal{A}. \tag{1.22}$$

The positivity of energy is the positivity (which we discussed above) of the operator H in the representation associated to the vacuum states.

In order to get "numbers" one bypasses the analytic difficulties of the above framework by working with the Feynman path integral formalism and in perturbation theory. It is in the latter framework that the problem of renormalization manifests itself in the simplest manner.

2.4. The simplest example.

Let us consider the simplest field theory, namely the free massive scalar field ϕ on the space-time

$$X = \mathbb{R} \times S^1$$

with the Lorentz metric.

The scalar field ϕ is a real-valued function on X governed by the Lagrangian

$$\mathcal{L}(\phi) = \frac{1}{2}\left((\partial_0\phi)^2 - (\partial_1\phi)^2 - m^2\,\phi^2\right)$$

where $\partial_0 = \partial/\partial t$ is the time derivative ($t \in \mathbb{R}$) and $\partial_1 = \partial/\partial x$ is the spatial derivative, where the space is here compact and one-dimensional. The action functional is given by

$$S(\phi) = \int \mathcal{L}(\phi)\,dx\,dt = \int L(t)\,dt$$

$$L(t) = \int_{S^1}\left(\frac{1}{2}\left(\dot{\phi}\right)^2 - \frac{1}{2}(\partial_1\,\phi)^2 - \frac{m^2}{2}\,\phi^2\right)\;dx.$$

Thus, at the classical level, one is dealing with a mechanical system with infinitely many degrees of freedom, whose configuration space C is the space of real-valued functions on S^1. The Hamiltonian of this classical mechanical system is the functional on the cotangent space T^*C given by

$$H(\phi,\pi) = \frac{1}{2}\int_{S^1}\left(\pi(x)^2 + (\partial\phi(x))^2 + m^2\phi^2(x)\right)\;dx,$$

where one uses the *linear* structure of C to identify T^*C with $C \times C^*$, and where one views the field π as an element of C^* in the form

$$\phi \mapsto \int_{S^1}\phi(x)\,\pi(x)\,dx \in \mathbb{C}.$$

As a classical mechanical system the above is the same as a countable collection of uncoupled harmonic oscillators. Indeed, one can take as coordinates in C (resp. C^*) the Fourier components $\phi_k = \int_{S^1}\phi(x)\;e^{-ikx}\;dx$ (resp. π_k), which are subject only to the reality condition

$$\phi_{-k} = \overline{\phi}_k \quad \forall k \in \mathbb{Z} \quad (\text{resp. } \pi_{-k} = \overline{\pi}_k).$$

Thus, both spaces C and T^*C are infinite products of finite-dimensional spaces and the Hamiltonian H is an infinite sum

$$H = \sum_{k\in\mathbb{Z}}\frac{1}{2}\left(\pi_k\overline{\pi}_k + (k^2+m^2)\phi_k\overline{\phi}_k\right).$$

The quantization of a single harmonic oscillator, say a system with configuration space $\mathbb{R}$ and Hamiltonian $\frac{1}{2}(p^2+\omega^2q^2)$, is given by the simple rule

$$[p,q] = -i\,\hbar\,.$$

This *quantizes* the values of the energy, as one sees from the equality (up to a shift of $\frac{\hbar\omega}{2}$)

$$H = \frac{1}{2}(p^2+\omega^2q^2) = \hbar\,\omega\,a^*a$$

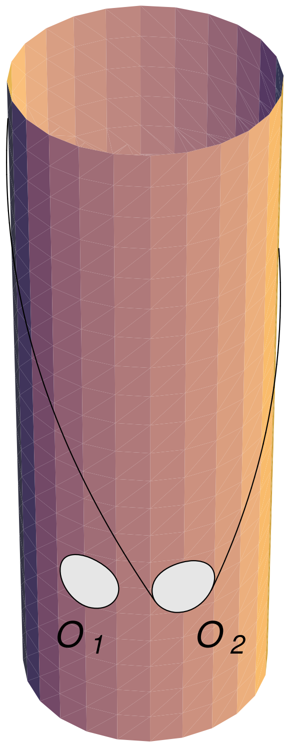


FIGURE 1. Space-like separated regions $\mathcal{O}_1$ and $\mathcal{O}_2$ in space-time cylinder

where the *annihilation* operator a and its adjoint, the *creation* operator a^*, obey the commutation relation

$$[a, a^*] = 1.$$

The terminology here comes from Dirac's theory of emission and absorption of photons [**119**]. These two equations completely describe the quantum system. Its representation in $L^2(\mathbb{R})$, given by the equality

$$a = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial q} + q\right)$$

is, up to unitary equivalence, its only irreducible representation. The (unique up to phase) normalized vector Ω such that $a\Omega = 0$ is called the vacuum vector.

The reality condition $\phi_{-k} = \overline{\phi}_k$ shows that for $k > 0$, the pair $\{-k, k\}$ yields a pair of harmonic oscillators, whose quantization yields a pair of creation operators a_k^*, a_{-k}^*. The algebra of observables of the quantized field has the following presentation. It is generated by a_k and a_k^*, $k \in \mathbb{Z}$, with relations

$$[a_k, a_k^*] = 1 \qquad \forall k \in \mathbb{Z}$$

$$[a_k, a_\ell] = 0 \ , \ [a_k, a_\ell^*] = 0 \qquad \forall k \neq \ell.$$

The Hamiltonian is given by the derivation corresponding to the formal sum

$$H_b = \sum_{k \in \mathbb{Z}} \hbar \omega_k a_k^* a_k \quad \text{with} \quad \omega_k = \sqrt{k^2 + m^2}. \tag{1.23}$$

The "vacuum representation" which corresponds to the vacuum state, which is the infinite tensor product of the vacuum states, is given by the Hilbert space

$$\mathcal{H}_b = \bigotimes_{k\in\mathbb{Z}} (\mathcal{H}_k, \Omega_k)$$

and the tensor product representation of the algebra.

Equivalently, one can describe this infinite tensor product as the Hilbert space $L^2(C, d\mu)$ where μ is a Gaussian measure on C (cf. [**145**]).

The spectrum of the quantum field Hamiltonian H_b acting in $\mathcal{H}_b$ is now

$$\text{Spec } H_b = \left\{\sum n_k \hbar\omega_k \mid n_k \in \mathbb{N}\right\}.$$

Quantum field theory reconciles positivity of energy (i.e. the fact that H_b is a *positive* operator in $\mathcal{H}_b$) with *causality*, which means that we have commutation at space-like separation for functions of the quantum field $\tilde{\phi}(t,x) = e^{it\cdot H_b}\, \tilde{\phi}(0,x)\, e^{-itH_b}$, with

$$\tilde{\phi}(0,x) = \sum_{k\in\mathbb{Z}} \left(a_k\, e^{ikx} + a_k^*\, e^{-ikx}\right)(2\omega_k)^{-1/2}.$$

This commutation of $\tilde{\phi}(t,x)$ with $\tilde{\phi}(s,y)$ for (t,x) and (s,y) space-like separated in the space-time $X = \mathbb{R}\times S^1$ is easy to check directly. First, the operators $\tilde{\phi}(f) = \int f(x)\, \tilde{\phi}(0,x)\, dx$, $f \in C^\infty(S^1)$, commute with each other, using $\omega_k = \omega_{-k}$. Next, with σ_t the automorphisms given by time evolution

$$\sigma_t = e^{itH_b}\,(\cdot)\, e^{-itH_b},$$

one computes $[\sigma_t\tilde{\phi}(f), \tilde{\phi}(g)]$, for $f, g \in C^\infty(S^1)$, and finds a scalar multiple of the identity, where the scalar is

$$\int c(x,y,t) f(x) g(y)\, dxdy,$$

where

$$c(x,y,t) = \sum_k e^{-ik(x-y)}\left(e^{-i\omega_k t} - e^{i\omega_k t}\right)\omega_k^{-1} = c(x-y,t)$$

with c satisfying the Klein-Gordon equation

$$\left(\partial_0^2 - \partial_1^2 + m^2\right) c = 0$$

with initial conditions $c(x,0) = 0$ and $\frac{\partial}{\partial t}\, c(x,0) = \lambda\, \delta_0(x)$ with $\lambda \neq 0$.

It then follows from elementary properties of the wave equation that $c(x,t)$ vanishes if (x,t) is space-like separated from $(0,0)$, hence that the quantum field $\tilde{\phi}$ satisfies causality. The classification of vacuum states is simple in this free case and one has the following result (cf. [**182**]).

a) The representation in $\mathcal{H}_b$ is the unique vacuum representation. The quantum fields $\tilde{\phi}$ are represented by operator-valued distributions.

b) Let $\mathcal{O}$ be a local region (i.e. a bounded open set) in X and $\mathcal{U}(\mathcal{O})$ be the von Neumann algebra in $\mathcal{H}_b$ generated by bounded functions of the $\tilde{\phi}(f)$ with $\mathrm{Supp}(f) \subset \mathcal{O}$. Then, when $\mathcal{O}_1$ and $\mathcal{O}_2$ are space-like separated, one has

$$\mathcal{U}(\mathcal{O}_1) \subset \mathcal{U}(\mathcal{O}_2)',$$

where $\mathcal{U}(\mathcal{O}_2)'$ is the commutant of the von Neumann algebra $\mathcal{U}(\mathcal{O}_2)$.

c) For all $\beta > 0$ one has $\mathrm{Tr}(e^{-\beta H_b}) < \infty$ and the map

$$a \mapsto \psi_\beta(a) = Z^{-1}\,\mathrm{Tr}(a\, e^{-\beta H_b}), \quad \text{with} \;\; Z = \mathrm{Tr}(e^{-\beta H_b}),$$

defines a state on the algebra of observables, which fulfills the KMS_β condition relative to σ_t. (See Chapter 3 below for a detailed discussion of the KMS_β condition.)

The von Neumann algebras $\mathcal{U}(\mathcal{O})$ of *local* observables are an essential part of the algebraic formulation of quantum field theory (cf. [**182**]). The finiteness of the trace $\mathrm{Tr}(e^{-\beta H_b}) < \infty$ follows from the convergence of the infinite product

$$\mathrm{Tr}(e^{-\beta H_b}) = \prod_{k\in\mathbb{Z}} \frac{1}{1 - e^{-\beta\,\omega_k}}$$

while the free energy $\log Z$ is given by

$$\log Z = -\sum_{k\in\mathbb{Z}} \log(1 - e^{-\beta\,\omega_k})\,. \tag{1.24}$$

2.5. Green's functions.

Once a vacuum state has been constructed, one way to describe the corresponding quantum field theory $\tilde{\phi}(x)$ is by means of the time-ordered Green's functions

$$G_N(x_1, \dots, x_N) = \langle\, 0\,| T\, \tilde{\phi}(x_1) \dots \tilde{\phi}(x_N)|\, 0\,\rangle, \tag{1.25}$$

where 0 is the vacuum state, and the time-ordering symbol T means that the $\tilde{\phi}(x_j)$'s are written in order of increasing time from right to left. Typically, from the Green's functions one can recover all the experimentally relevant data, such as the cross sections of the particle collisions.

In the functional integral formalism as in (1.15), the Green's functions can be computed by Feynman's formula as an "integral over histories"

$$G_N(x_1, \dots, x_N) = \mathcal{N} \int e^{i\,\frac{S(\phi)}{\hbar}}\; \phi(x_1)\cdots\phi(x_N)\,\mathcal{D}[\phi], \tag{1.26}$$

where the factor $\mathcal{N}$ ensures the normalization of the vacuum state

$$\langle\, 0 \mid 0\,\rangle = 1\,. \tag{1.27}$$

This means that, if one could ignore the issue of renormalization, which we will discuss later, then the functional integral (1.26) would be easy to compute in perturbation theory, i.e. by treating the term $\mathcal{L}_{\mathrm{int}}$ in (1.12) as a perturbation of (1.14). The action functional (1.11) correspondingly splits as the sum of two terms

$$S(\phi) = S_0(\phi) + S_{\mathrm{int}}(\phi), \tag{1.28}$$

where the free action S_0 generates an "imaginary Gaussian" measure

$$d\mu = \exp\left(i\, S_0(\phi)\right)\mathcal{D}[\phi], \tag{1.29}$$

where we have set $\hbar = 1$.

The series expansion of the Green's functions can then be written in the form

$$G_N(x_1,\dots,x_N) = \left(\sum_{n=0}^\infty \frac{i^n}{n!}\int \phi(x_1)\cdots\phi(x_N)\, S_{\rm int}(\phi)^n\, d\mu\right) \cdot \left(\sum_{n=0}^\infty \frac{i^n}{n!}\int S_{\rm int}(\phi)^n\, d\mu\right)^{-1} \tag{1.30}$$

or equivalently in terms of the Green's functions for the free field $\tilde\phi_F$ as the Gell-Mann–Low formula

$$G_N(x_1,\dots,x_N) = \left(\sum_{n=0}^\infty \frac{i^n}{n!}\int\langle 0\,|T\tilde\phi_F(x_1)\cdots\tilde\phi_F(x_N)\,\prod_j \mathcal{L}_{\rm int}(y_j)\,|\,0\,\rangle\prod_j dy_j\right) \cdot\left(\sum_{n=0}^\infty \frac{i^n}{n!}\int\langle 0\,|T\,\prod_j \mathcal{L}_{\rm int}(y_j)\,|\,0\,\rangle\prod_j dy_j\right)^{-1}. \tag{1.31}$$

2.6. Wick rotation and Euclidean Green's functions.

Even for the simplest case of the two-point Green's function of the free field $\tilde\phi_F$ some care is required in order to get the right answer. Indeed, in terms of (1.30) the only non-zero term is for $n = 0$ which gives, with the normalization factor $\mathcal{N}_0$ of the free Gaussian,

$$G_2^F(x,y) = \mathcal{N}_0 \int \phi(x)\phi(y)\, \exp\left(i\, S_0(\phi)\right)\mathcal{D}[\phi]. \tag{1.32}$$

As a function of the classical field ϕ, the expression $\phi(x)\phi(y)$ only depends upon the two-dimensional projection of ϕ but one cannot yet use the factorization property of the Gaussian integral since $S_0(\phi)$ involves the derivatives of ϕ. In fact $S_0(\phi)$ is diagonal in terms of the Fourier transform $\hat\phi(p)$,

$$\hat\phi(p) = \int \phi(x)\, e^{-ip\cdot x}\, d^D x, \tag{1.33}$$

where it becomes

$$S_0(\phi) = (2\pi)^{-D}\int \frac{1}{2}(p^2 - m^2)\,\hat\phi(p)\hat\phi(-p)\, d^D p. \tag{1.34}$$

The reality condition for $\phi(x)$ shows that the $\hat\phi(p)$ are not independent variables but fulfill the condition

$$\hat\phi(-p) = \overline{\hat\phi(p)}. \tag{1.35}$$

Then one uses the Fourier inversion formula

$$\phi(x) = (2\pi)^{-D}\int \hat\phi(p)\, e^{ip\cdot x}\, d^D p \tag{1.36}$$

and applies integration by parts under the Gaussian to get

$$\mathcal{N}_0 \int \hat\phi(p_1)\,\hat\phi(p_2)\,\exp\left(i\, S_0(\phi)\right)\mathcal{D}[\phi] = i\,(2\pi)^D\,\delta(p_1+p_2)\,(p_1^2 - m^2)^{-1}. \tag{1.37}$$

At the formal level, this integration by parts is obtained from the following simple manipulations. One lets V be a finite-dimensional real vector space with $Q \in V^*\otimes V^*$

a non-degenerate quadratic form. The inverse of Q is simply obtained from the canonical identification $V^* \otimes V^* \sim \mathrm{Hom}(V, V^*)$ as

$$\mathrm{Hom}(V, V^*) \ni Q \mapsto Q^{-1} \in \mathrm{Hom}(V^*, V).$$

The invertibility and the symmetry of Q show that, for any linear form $L \in V^*$ on V, one has

$$\partial_{Q^{-1}(L)} \frac{1}{2} Q = L\,.$$

The integration by parts formula is then obtained from

$$\int P(X)\, L(X)\, \exp\left(-\frac{Q(X)}{2}\right) \mathcal{D}[X] =$$
$$-\int P(X)\, \partial_{Q^{-1}(L)} \left(\exp\left(-\frac{Q(X)}{2}\right)\right) \mathcal{D}[X] =$$
$$\int \partial_{Q^{-1}(L)}(P(X))\, \exp\left(-\frac{Q(X)}{2}\right) \mathcal{D}[X].$$

This gives

$$\begin{aligned} &\textstyle\int P(X)\, L(X)\, \exp\left(-\frac{Q(X)}{2}\right) \mathcal{D}[X] = \\ &\textstyle\int \partial_{Q^{-1}(L)}(P(X)) \exp\left(-\frac{Q(X)}{2}\right) \mathcal{D}[X]. \end{aligned} \tag{1.38}$$

Thus, (1.37) reduces the computation of the two-point Green's function (1.32) to the integral

$$G_2^F(x, y) = i\,(2\pi)^{-D} \int \frac{e^{\pm i p\cdot(x-y)}}{p^2 - m^2}\, d^D p. \tag{1.39}$$

The presence of the oscillatory term $e^{-ip\cdot(x-y)}$ takes good care of the issue of convergence for large p. The reason why care is needed is the singularity at $p^2 = m^2$ which makes the meaning of (1.39) ambiguous.
The right answer is easy to compute for the explicit example of subsection 2.4, where one gets by direct computation

$$\langle 0 |\, \phi_F(t, x) \phi_F(s, y) \,| 0 \rangle = \sum_{k\in\mathbb{Z}} (2\omega_k)^{-1}\, e^{i(k(x-y)-(t-s)\omega_k)}, \tag{1.40}$$

where, as in (1.23), the frequency ω_k is the *positive* square root of $k^2 + m^2$.

LEMMA 1.2. *Let $u > 0$ and $\omega > 0$ be positive real numbers. Then one has*

$$\lim_{\epsilon\to 0+} \int_{\mathbb{R}} \frac{e^{\pm i p u}}{p^2 - \omega^2 + i\epsilon}\, dp = \frac{\pi}{i\,\omega}\, e^{-iu\omega}\,.$$

PROOF. The integral

$$F(\omega) = \int_{\mathbb{R}} \frac{e^{ipu}}{p^2 - \omega^2}\, dp$$

defines a holomorphic function of ω in the half-plane $\Im(\omega) < 0$. For $a > 0$ one has $F(-ia) = \frac{\pi}{a}\, e^{-ua}$, hence

$$F(\omega) = \frac{\pi}{i\,\omega}\, e^{-iu\omega}\,, \quad \forall \omega \in \mathbb{C} \quad \text{with} \quad \Im(\omega) < 0.$$

Using this as a formula for the square root of $\omega^2 - i\epsilon$ near ω one gets the desired result. $\square$

This makes it possible to rewrite (1.40) using Dyson's time-ordered Green's function, in the form

$$\langle 0|\, T\, \phi_F(t,x)\phi_F(s,y)\,|\,0\,\rangle = \frac{i}{2\pi}\sum_{k\in\mathbb{Z}}\int_{\mathbb{R}} \frac{e^{i(k(x-y)-k_0(t-s))}}{k_0^2-k^2-m^2+i\epsilon}\, dk_0, \tag{1.41}$$

where the limit for $\epsilon \to 0$ is understood. The role of the time ordering T is simply in order to ensure $(t-s) \geq 0$. Thus, more generally, the prescription due to Feynman for the removal of the singularity (or on how to go around the pole) in the integral (1.39) consists of defining it as

$$G_2^F(x,y) = i\,(2\pi)^{-D}\int \frac{e^{\pm ip\cdot(x-y)}}{p^2-m^2+i\epsilon}\, d^D p. \tag{1.42}$$

This prescription embodies the physical conditions fulfilled by the free field theory and makes it possible to start the perturbation theory. It is important to be able to formulate it at a more conceptual level and this is done by the Wick rotation to imaginary time. The Green's functions (1.25) then appear as the distributional boundary values of the Schwinger functions $S_N(x_1,\ldots,x_N)$. These are complex analytic functions of the complexified time parameters z_j of $x_j = (z_j, v_j)$. They are defined in the cone

$$\Im\, z_1 < \Im\, z_2 < \cdots < \Im\, z_N. \tag{1.43}$$

The relevance of the Schwinger functions $S_N(x_1,\ldots,x_N)$ comes from the positivity of the Hamiltonian H and the formal equality

$$S_N(x_1,\ldots,x_N) = \langle\, 0\,|\,\tilde\phi(v_1)\, e^{i\,(z_2-z_1)\,H}\,\cdots\,\tilde\phi(v_{N-1})\, e^{i\,(z_N-z_{N-1})\,H}\,\tilde\phi(v_N)|\,0\,\rangle.$$

In particular, for $N=2$, the ambiguity in the definition of the Green's function of the free field is uniquely resolved in terms of the corresponding Schwinger function. The existence of the two-point Schwinger function

$$\langle\, 0\,|\,\tilde\phi(v_1)\, e^{i\,z\,H}\,\tilde\phi(v_2)|\,0\,\rangle\,, \quad \Im\, z \geq 0 \tag{1.44}$$

is a limit case (at zero temperature) of the Kubo–Martin–Schwinger (KMS) condition which plays a central role in Chapter 3 of this book.

To perform the computation one works directly with purely imaginary time. Thus, one lets $\tau = it$ and one expresses the Schwinger functions in the form

$$S_N(x_1,\ldots,x_N) = \mathcal{N}\int \phi_E(x_1)\ldots\phi_E(x_N)\, e^{-\frac{S_E(\phi_E)}{\hbar}}\,\mathcal{D}[\phi_E], \tag{1.45}$$

where $\mathcal{N}$ is the normalization factor

$$\mathcal{N}^{-1} = \int e^{-\frac{S_E(\phi_E)}{\hbar}}\,\mathcal{D}[\phi_E]\,,$$

and where the Euclidean action functional

$$S_E\,(\phi_E) = \int \mathcal{L}_E\,(\phi_E)\, d^D x \tag{1.46}$$

is given by the Lagrangian density

$$\mathcal{L}_E\,(\phi_E) = \frac{1}{2}(\partial\phi_E)^2 + \frac{m^2}{2}\,\phi_E^2 + \mathcal{L}_{\rm int}(\phi_E)\,. \tag{1.47}$$

The integration variables ϕ_E are *classical Euclidean fields* and one uses the (positive) Euclidean metric $g^{\mu\nu} = \delta^{\mu\nu}$ to compute $(\partial\phi)^2 = g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$. The Euclidean functional integral

$$\langle \mathcal{O} \rangle_E = \mathcal{N} \int \mathcal{O}(\phi_E)\, e^{-\frac{S_E(\phi_E)}{\hbar}}\, \mathcal{D}[\phi_E] \tag{1.48}$$

for an observable $\mathcal{O}(\phi_E)$ (a function of the classical Euclidean fields) is identical to the expectation value of a classical statistical mechanical system and this gives a far-reaching correspondence between Euclidean field theory and statistical mechanics ([**145**]).

Notice that the change of sign of the potential terms of the Lagrangian in the transition from Minkowski space-time to Euclidean space-time is the same as in the transition (1.7) from the Lagrangian to the Hamiltonian. In particular, one keeps track of the basic *positivity* properties of the latter, the measure $d\mu$ of (1.29) becomes a true Gaussian measure after the Wick rotation to Euclidean signature, and the singularity of the propagator (cf. (1.39)) has now disappeared.

The series expansion of the Euclidean Green's functions takes the form

$$\begin{aligned} &S_N(x_1, \ldots, x_N) = \\ &\left(\textstyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \phi_E(x_1) \cdots \phi_E(x_N)\, S_{\rm int}(\phi_E)^n\, d\mu\right) \\ &\cdot \left(\textstyle\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int S_{\rm int}(\phi_E)^n\, d\mu\right)^{-1} . \end{aligned} \tag{1.49}$$

3. Feynman diagrams

The various terms

$$\int \phi(x_1) \ldots \phi(x_N)\, (S_{\rm int}(\phi))^n\, d\mu \tag{1.50}$$

of the expansion (1.30) are integrals of polynomials with respect to a Gaussian measure $d\mu$. When these are computed using integration by parts, the process generates a large number of terms. The combinatorial data labeling these terms are conveniently encoded by graphs Γ, known as Feynman graphs. Each graph determines the terms that appear in the calculation of the corresponding integral.

The class of graphs that label terms in this procedure depends on the nature of the terms in the Lagrangian density of the theory. For example, the possible valences of vertices in the graphs correspond to the monomials in the Lagrangian density: the quadratic terms give rise to vertices of valence two, while, for instance, the presence of a term $g\phi^3$ in the interaction Lagrangian $\mathcal{L}_{\rm int}$ means that there will be vertices of valence three in the graphs.

The physical intuition behind Feynman diagrams is that each graph represents a possible event where particles are exchanged, created and annihilated. For example, the graph depicted in Figure 2 shows an electron releasing a photon, which in turn creates a pair consisting of a particle and antiparticle (an electron and a positron). These are then annihilated releasing a photon that is finally absorbed by an electron. Diagrams of increasing complexity enumerate more complicated possible interactions.

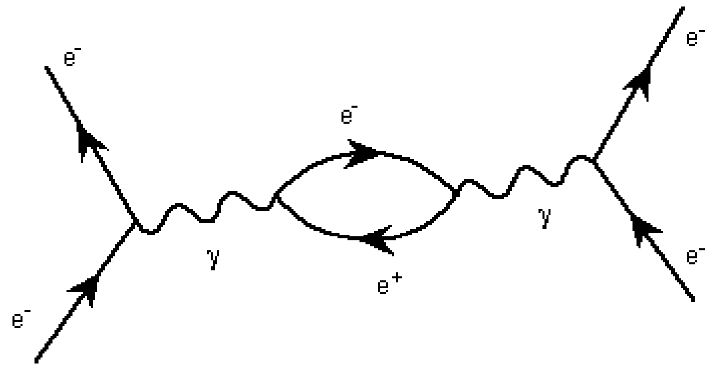


FIGURE 2. An example of a Feynman diagram

The particles that are involved in the intermediate processes are "virtual" particles which differ notably from real particles in that their energy-momentum vectors p no longer satisfy the "on-shell" condition

$$p^2 = m^2\,, \quad p_0 \geq 0 \tag{1.51}$$

and the integration process involves all "virtual" particles including those of negative energy.

3.1. The simplest case.

In order to understand how Feynman graphs label the various terms of the perturbative expansion, let us take the simplest interaction term, namely a monomial of the form ϕ^3, so that the Lagrangian is given by

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\,\phi^2 - \frac{g}{3!}\,\phi^3 \tag{1.52}$$

where g is the coupling constant.

The first non-trivial correction to the Euclidean two-point function comes from the term in (1.49) of the form

$$\int \phi_E(x_1)\,\phi_E(x_2)\,S_{\rm int}(\phi_E)^2\,d\mu, \tag{1.53}$$

where, by construction, we have

$$\begin{aligned} S_{\rm int}(\phi_E) = &\ \tfrac{g}{3!}\int \phi_E^3(x)\,d^Dx \\ = &\ (2\pi)^{-2D}\tfrac{g}{3!}\int \hat\phi(k_1)\hat\phi(k_2)\hat\phi(k_3)\delta(k_1+k_2+k_3)\prod d^Dk_j. \end{aligned}$$

The simple mnemonic, in order to keep track of the powers of 2π, is to attach a factor of $(2\pi)^{-D}$ to each momentum integration d^Dk and a factor of $(2\pi)^D$ to each δ-function. Using

$$\phi_E(x_j) = (2\pi)^{-D}\int \hat\phi(p)\,e^{ipx_j}\,d^Dp$$

the problem is then reduced, up to an overall factor and to the product

$$(2\pi)^{2D}\delta(k_1+k_2+k_3)\delta(q_1+q_2+q_3),$$

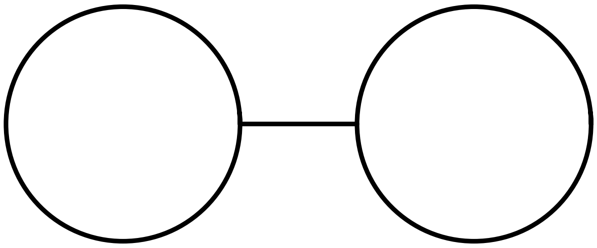


FIGURE 3. Vacuum bubble

to computing the integral

$$\int \hat{\phi}(p_1)\hat{\phi}(p_2)\hat{\phi}(k_1)\hat{\phi}(k_2)\hat{\phi}(k_3)\hat{\phi}(q_1)\hat{\phi}(q_2)\hat{\phi}(q_3)\, d\mu, \tag{1.54}$$

which is the integral of a polynomial of degree 8 under a Gaussian. To compute it, one just needs to apply (1.38) four times, contracting pairwise the factors of the monomial. This generates $7 \times 5 \times 3 = 105$ terms. Each of the contractions takes place between pairs of factors of the form $\hat{\phi}(\ell_1)\,\hat{\phi}(\ell_2)$ and delivers a delta-function

$$(2\pi)^D\,\delta(\ell_1 + \ell_2)$$

in the sum of the momenta ℓ_j involved, and a "propagator", i.e. a factor of

$$\frac{1}{\ell^2 + m^2}, \tag{1.55}$$

where $\ell = \ell_j$ and the choice of j is irrelevant since the propagator is an even function of the momentum.
The terms where $\hat{\phi}(p)$ is paired with $\hat{\phi}(-p)$ are not relevant, since they get canceled by the denominator of (1.49). It is nevertheless worthwhile to see what they look like. This means computing the degree-six expression

$$(2\pi)^{2D}\delta(k_1 + k_2 + k_3)\delta(q_1 + q_2 + q_3) \int \hat{\phi}(k_1)\,\hat{\phi}(k_2)\,\hat{\phi}(k_3)\,\hat{\phi}(q_1)\,\hat{\phi}(q_2)\,\hat{\phi}(q_3)\, d\mu.$$

In any of the 15 possible pairings, at least one of the $\hat{\phi}(k)$ will get paired with one of the $\hat{\phi}(q)$. There are then two possible cases:

- Only one of the $\hat{\phi}(k)$ is paired with one of the $\hat{\phi}(q)$;
- The three $\hat{\phi}(k)$ are paired with the three $\hat{\phi}(q)$.

An example of the first case is given by the pairings

$$k_1 \longleftrightarrow k_2\,, \quad k_3 \longleftrightarrow q_1\,, \quad q_2 \longleftrightarrow q_3.$$

There are 9 such pairings. An example of the second case (which yields 6 pairings) is given by

$$k_1 \longleftrightarrow q_1\,, \quad k_2 \longleftrightarrow q_2\,, \quad k_3 \longleftrightarrow q_3.$$

These two possibilities are represented pictorially by the *vacuum bubble* diagrams of Figures 3, 4.

In the other terms, the factors of $\hat{\phi}(p)$ and $\hat{\phi}(-p)$ get paired with the terms of the monomial $\hat{\phi}(k_1)\,\hat{\phi}(k_2)\,\hat{\phi}(k_3)\,\hat{\phi}(q_1)\,\hat{\phi}(q_2)\,\hat{\phi}(q_3)$. Thus, there are again two distinct cases:

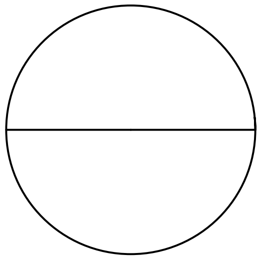


FIGURE 4. Vacuum bubble

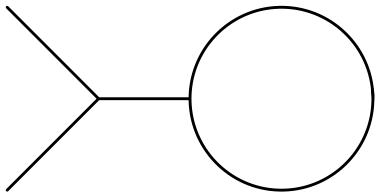


FIGURE 5. Tadpole

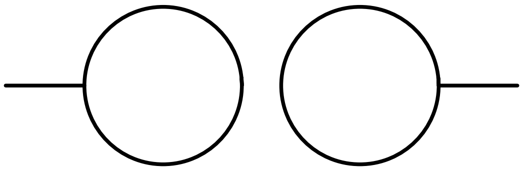


FIGURE 6. Double tadpole

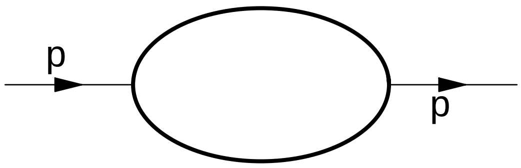


FIGURE 7. Self-energy graph

- The terms $\hat{\phi}(p)$ and $\hat{\phi}(-p)$ get paired with terms of the same group;
- The terms $\hat{\phi}(p)$ and $\hat{\phi}(-p)$ get paired with terms of two different groups (i.e. with a k and a q).

These cases are represented pictorially by the diagrams of Figure 5 for the first case, and Figure 6 and Figure 7 for the second case. The latter, in fact, splits into two subcases according to the pairings of the remaining terms in the two groups.

The presence of the δ-functions (i.e. momentum conservation) shows that in all the tadpole graphs the momentum k flowing through the single line is zero (i.e. there is an overall factor of $\delta(k)$). The contributions of all these diagrams can be eliminated by the following a priori requirement on the vacuum state:

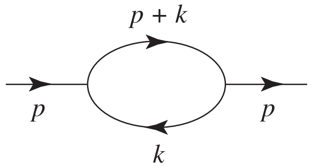
$$\langle 0 \,|\, \phi(x) \,|\, 0 \rangle = 0. \tag{1.56}$$

This condition is satisfied in the free field case and one requires that it continues to hold in the perturbative case (tadpole vanishing condition).

Thus, we focus on the only remaining graph, that is, the self-energy graph of Figure 7. It corresponds to one of the 36 pairings of the form

$$p_1 \longleftrightarrow k_1\,, \quad p_2 \longleftrightarrow q_1\,, \quad k_2 \longleftrightarrow q_2\,, \quad k_3 \longleftrightarrow q_3.$$

Each of the edges delivers a propagator. Thus, the external lines (called *external legs* of the graph) each give rise to a term in $(p^2+m^2)^{-1}$. This means that we get an overall factor of $(p^2+m^2)^{-2}$. Notice then that the δ-functions occurring from the contractions and from the overall factor $\delta(k_1+k_2+k_3)\delta(q_1+q_2+q_3)$ do not uniquely determine the momenta k_j and q_j from the value of p. Indeed, there is one law of conservation of momentum per vertex of the graph and one free momentum variable for each internal line. However, the obtained linear relations are not independent. They leave one momentum variable k undetermined, so that the flow of momenta through the diagram is given by



Including the δ-functions and the coupling constant, the contribution of the self-energy graph is then given by

$$\frac{1}{2} g^2\, \delta(p_1+p_2) \frac{1}{p_1^2+m^2} \frac{1}{p_2^2+m^2} \int \frac{1}{k^2+m^2} \frac{1}{(p_1+k)^2+m^2}\, d^D k. \tag{1.57}$$

Notice here that there are 36 pairings giving the same graph and the factor of $(\frac{1}{3!})^2$ in front of g^2 cancels this numerical factor.

Thus, we see that in dimension $D = 4$ the problem of divergences shows up when one wants to integrate the contributions for the different values of k as required in computing (1.53). In the expression (1.57) large values of the Euclidean momentum k contribute disproportionate amounts to the self-energy and if one integrates only over momenta k with $\|k\| \leq \Lambda$ where $\Lambda > 0$ is a *cutoff* parameter, one obtains

$$\int_{\|k\| \leq \Lambda} \frac{1}{k^2+m^2} \frac{1}{(p_1+k)^2+m^2}\, d^4 k \sim 2\pi^2 \log \Lambda, \tag{1.58}$$

which diverges when the cutoff is removed, i.e. for $\Lambda \to \infty$. It would then appear at first sight that the theory does not make sense, since it delivers divergent answers. This type of problem arose in the early 1930's, when Oppenheimer, starting from Dirac's theory, computed the second-order correction corresponding to the emission and absorption of a virtual photon by an electron, as in Figure 8. and concluded that self-energy effects generate an infinite displacement of the spectral lines. The physics resolution of this type of problems is called renormalization.

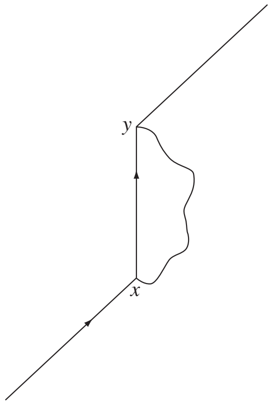


FIGURE 8. Self-energy graph in QED

3.2. The origins of renormalization.

The idea of renormalization goes back to the mid-nineteenth century, long before the advent of quantum field theory. At the time, the problem that motivated Green to introduce a form of renormalization was the motion of bodies in fluids [**150**]. It was observed that, due to the interaction with the surrounding field of water, a moving object still obeys the usual physical rules, like Newton's law $F = ma$, but with the mass m no longer given by the "bare" mass m_0 of the body, but by a corrected "physical" mass $m = m_0 + \delta m$.

More precisely, this statement holds for a spherical object B immersed in a perfect incompressible fluid with constant density ρ. When B is moving at a speed v, say along the z-axis, it imparts a speed distribution X to the fluid surrounding B which can be computed, in the moving frame of B, as the gradient vector field of the solution h of the Laplace equation in the complement of B whose gradient is tangent to ∂B. Thus one has

$$X = \nabla h, \quad \text{with } \Delta h = 0.$$

This solution is given in terms of the first spherical harmonics and is of the form

$$h(x, y, z) = \frac{v}{2}\,(r^{-3} + 2)\,z, \quad \text{where } r^2 = x^2 + y^2 + z^2,$$

in the moving frame where B is the ball of radius one centered at 0. Thus, with respect to the fluid, the vector field is $Y = \nabla h - (0, 0, v)$ and its kinetic energy density is easily computed to be

$$E(x, y, z) = \frac{1}{2}\,\rho \|Y\|^2 = \frac{v^2}{8}\,\frac{x^2 + y^2 + 4z^2}{(x^2 + y^2 + z^2)^4}\,\rho.$$

The energy of the fluid disturbance is then the integral

$$\int_{B^c} E(x, y, z)\,dxdydz = \lambda\rho\, v^2$$

in the complement B^c of B. An easy direct computation shows that one has

$$\lambda\rho = \frac{1}{4}\,M\,, \quad \lambda\rho\, v^2 = \frac{1}{4}\,M v^2,$$

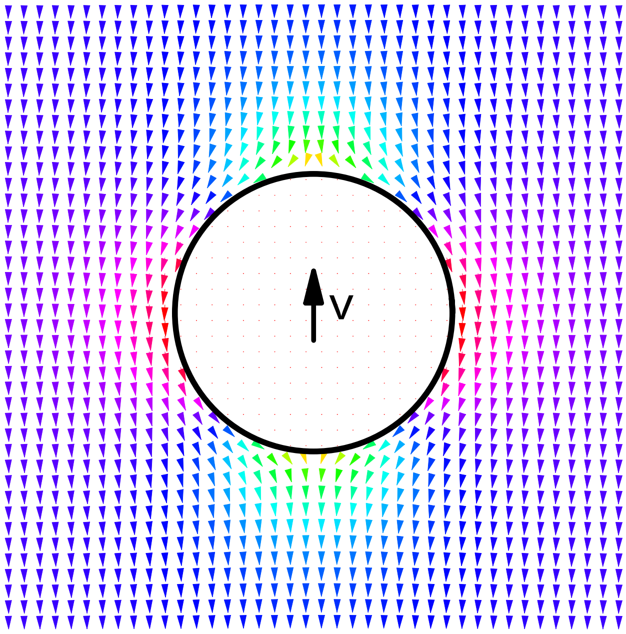


FIGURE 9. Self-energy of an immersed ball

where M is the mass of the fluid contained in the volume of the ball B. As shown by Green [**150**], this implies that the inertial mass m that appears in Newton's law of motion for the ball B undergoes, by the mere presence of the surrounding fluid, a "renormalization"

$$m \mapsto m + \delta m = m + \frac{M}{2}.$$

The striking fact is that this effect is felt even at zero speed and it modifies the initial acceleration of the moving ball.

An amusing example of what this means in very concrete terms is given in Sidney Coleman's course on quantum field theory [**59**]. Namely, one considers what will be the initial acceleration of a ping-pong ball immersed in water, subject to the usual Archimedes law. If one simply applies Newton's law with the bare mass of the ball, one finds an unrealistically high value for the acceleration (something of the order of $11\,g$), while if one corrects the value of the mass as observed by Green, where $\delta m = M/2$ is half of the mass of the water contained in the volume of the ball, then Archimedes law $-(M - m_0)g = ma$ yields a more realistic value of the acceleration of the order of $2\,g$. (The ball has a mass of $m_0 = 2.7$ grams and a diameter of 4 centimeters. (Readers not versed in the metric system can make a conversion to their favorite units.)

The additional inertial mass $\delta m = m - m_0$ accounts for the interaction of the body with the surrounding field of water. In this classical setup with bodies moving in fluids, clearly both m and m_0 are experimentally measurable. However, if the interaction could not be turned off by removing the body from the fluid, there would

Figure 10. Tree level counterterm graph

be no way to measure the bare mass m_0, and m would remain the only physically observable mass.

The analogy between hydrodynamics and electromagnetism led in steps, through the work of Thomson, Lorentz, Kramers, et al. (cf. [**125**]), , to the crucial distinction between the bare parameters, such as m_0, which enter the field-theoretic equations, and the observed parameters, such as the inertial mass m. Since it is not possible to isolate an electron from the surrounding electromagnetic field, in quantum field theory the bare parameters are not physical observables.

Around 1947, motivated by the experimental findings of spectroscopy about the fine structure of spectra, physicists were able to exploit the distinction between these two notions of mass (bare and observed), and similar distinctions for the charge and field strength, in order to eliminate the unwanted infinities associated to the pointwise nature of the electron. This marked the beginning of renormalization techniques in quantum field theory. We refer the interested reader to [**125**] for a detailed account of the historical developments of the subject.

In the example of §3.1 above, one uses the aforementioned distinction so as to cancel the second-order self-energy divergence in the following way. One takes the freedom to shift the bare mass parameter m^2 which appears in the Lagrangian by an amount δm^2 which depends on the cutoff parameter Λ. This results in a new term

$$\delta S(\phi_E) = -\frac{1}{2}\,\delta m^2(\Lambda) \int \phi_E^2(x)\, d^4x \tag{1.59}$$

in the Euclidean action. Considering this term as part of the interaction $S_{\rm int}$, one obtains a corresponding term in the perturbative expansion (1.49) of the Euclidean two-point function, of the form

$$-\int \phi_E(x_1)\,\phi_E(x_2)\,\delta S(\phi_E)\,d\mu. \tag{1.60}$$

Using (1.59), one obtains

$$-\delta S(\phi_E) = \frac{1}{2}\,\delta m^2(\Lambda)\,(2\pi)^{-4} \int \hat\phi(k_1)\,\hat\phi(k_2)\,\delta(k_1+k_2) \prod d^4k_j. \tag{1.61}$$

Thus, we get the additional term

$$\frac{1}{2}\,\delta m^2(\Lambda)\,(2\pi)^{-4}\,\delta(k_1+k_2)\,(\int \hat\phi(p_1)\,\hat\phi(p_2)\,\hat\phi(k_1)\,\hat\phi(k_2)\,d\mu) \prod d^4k_j, \tag{1.62}$$

which is the integral of a polynomial of degree 4 under a Gaussian. When we apply (1.38), the terms in which the p_j get paired together only contribute to the normalization factor and are canceled out. Thus, we just get the two terms in which p_1 is paired with one of the k_j. These are represented by the graph of Figure 10 and their values add up to

$$\delta m^2(\Lambda)\,(2\pi)^4\,\delta(p_1+p_2)\frac{1}{p_1^2+m^2}\,\frac{1}{p_2^2+m^2}. \tag{1.63}$$

This shows that the choice

$$\delta m^2(\Lambda) = -(2\pi)^{-4}\,\pi^2\,g^2\,\log\Lambda, \tag{1.64}$$

cancels the divergence coming from the second-order self-energy graph. In other words, one can take advantage of the unphysical nature of the bare parameters to ask them to conspire so that they cancel the divergences. This is a very tricky process since a counterterm such as (1.64) will now generate many other new terms in the perturbative expansion and it is not clear at all how to organize the elimination of the unwanted infinities. After clarifying what is meant by a Feynman graph (subsection 3.3) we describe the first step, which is to reduce the zoo of graphs that are physically relevant and this is done in terms of generating functions and the effective action (subsection 3.5).

3.3. Feynman graphs and rules.

As explained in §3.1 above, the various terms of the perturbative expansion of Euclidean Green's functions (1.49) are labeled by graphs which encode in a visual manner all the possible pairings in the integration by parts under the Gaussian. Thus, at the formal level, one has

$$S_N(x_1,\ldots,x_N) = \sum \int \frac{V(\Gamma)(p_1,\ldots,p_N)}{\sigma(\Gamma)} e^{i(x_1\cdot p_1+\cdots+x_N\cdot p_N)} \prod_j \frac{dp_j}{(2\pi)^D}, \tag{1.65}$$

where we need to give precise meaning to the terms

- Γ (Feynman graph)
- $V(\Gamma)(p_1,\ldots,p_N)$ (unrenormalized value of the graph)
- $\sigma(\Gamma)$ (symmetry factor).

Graphs: A Feynman graph Γ is given by a finite set $\Gamma^{(0)}$ of vertices and a finite set $\Gamma^{(1)}$ of oriented edges. (The oriented edges are also called lines. We are dealing here, for simplicity, with a single scalar field and do not have to distinguish among several propagators.) The sets of vertices and edges are endowed with a collection of maps ∂_j, for $j \in \{0,1\}$ and ι. Here

$$\partial_j : \Gamma^{(1)} \to \Gamma^{(0)} \cup \{1,2,\ldots,N\} \tag{1.66}$$

is the map that assigns to each line the two vertices (corresponding to the two cases $j \in \{0,1\}$) of its boundary. Let I denote the collection of monomials in the interaction Lagrangian $\mathcal{L}_{\rm int}$. Then the map ι is given by

$$\iota : \Gamma^{(0)} \to I. \tag{1.67}$$

It assigns to each vertex a monomial in the interaction Lagrangian $\mathcal{L}_{\rm int}$. One requires that the degree of the monomial is the valence of the vertex, i.e. that

$$\deg \iota(v) = \#\partial_0^{-1}(v) + \#\partial_1^{-1}(v)\,, \quad \forall v \in \Gamma^{(0)}. \tag{1.68}$$

One also requires that there are N external lines labeled by $\{1,2,\ldots,N\}$. The labelling is by the vertex in $\{1,2,\ldots,N\}$ and one assumes that

$$\#\partial_0^{-1}(v) + \#\partial_1^{-1}(v) = 1\,, \quad \forall v \in \{1,2,\ldots,N\}. \tag{1.69}$$

The set of external lines is

$$\Gamma^{(1)}_{\rm ext} = \cup_{j=0}^{1} \partial_j^{-1}\{1,2,\ldots,N\} \subset \Gamma^{(1)} \tag{1.70}$$

and its complement $\Gamma^{(1)}_{\rm int} \subset \Gamma^{(1)}$ is the set of internal lines.

The geometric realization of a graph is the one-dimensional space

$$|\Gamma| = \Gamma^{(1)} \times [0,1] \cup_\partial (\Gamma^{(0)} \cup \{1,2,\ldots,N\}) \tag{1.71}$$

obtained by gluing the endpoints of the lines using the maps ∂_j. The graph is *planar* when it can be represented by a planar picture. It can always be represented by a spatial picture, hence by a planar projection with a finite numbers of crossings (usually indicated by a broken line, so as not to be confused with vertices).

Graphs from pairings: We write the decomposition of the interaction Lagrangian $\mathcal{L}_{\rm int}$ into a sum of monomials as $\mathcal{L}_{\rm int} = \sum_{\mathcal{M}\in I} \mathcal{M}$. We then look at the contribution to the perturbative expansion (1.49) coming from terms of the form

$$\frac{(-1)^n}{n!} \int \hat\phi(p_1) \cdots \hat\phi(p_N) \prod_{j\in X} \mathcal{M}_j \, d\mu, \tag{1.72}$$

where X is a finite set mapping to I, by $X \ni j \mapsto \mathcal{M}_j \in I$. Each monomial of the form

$$\mathcal{M} = \frac{-z}{d!}\,\phi_E(x)^d$$

yields, as in §3.1, a product

$$z\,(2\pi)^D\,\delta(\sum k_i) \prod \hat\phi(k_i)\,\frac{dk_i}{(2\pi)^D}, \tag{1.73}$$

where we omit the integral in k_i.

The derivative interaction term $\frac{-z}{2}(\partial\phi_E(x))^2$ delivers a quadratic term

$$z\,(2\pi)^D\,\delta(k_1+k_2)k_1^2\hat\phi(k_1)\hat\phi(k_2)\frac{dk_1 dk_2}{(2\pi)^{2D}}. \tag{1.74}$$

Notice that in the original Lagrangian this term is not an interaction term, but later on, in the process of wave function renormalization, one needs to add such a term and view it as part of the interaction.

The terms (1.72) generate graphs Γ from the pairings in the integration by parts in

$$\int \hat\phi(p_1) \cdots \hat\phi(p_N) \prod_{j\in X} \prod_{i=1}^{d_j} \hat\phi(k_i(j))\, d\mu. \tag{1.75}$$

For each such graph Γ the set of vertices is $\Gamma^{(0)} = X$ and the map ι is the obvious one which, to each element $j \in X$, assigns the monomial $\mathcal{M}_j$. Thus, each monomial determines a vertex with a number of half-lines equal to its degree, and a pairing in the integration by parts (1.75) corresponds to a way of matching these half-lines. In (1.75) the $k_i(j)$ denote the momentum variables carried by the half-lines associated to the monomial $\mathcal{M}_j$. Thus, we see that each such pairing in (1.75) determines in this way a graph Γ, where the set $\Gamma^{(1)}$ of edges is the set of the coupled pairs. For each such pair we make an arbitrary choice of orientation. The map ∂ in (1.71) is obtained by taking the image of the two elements of the pair under the map

$$p_\ell \mapsto \ell \in \{1,2,\ldots,N\}, \quad k_i(j) \mapsto j \in X = \Gamma^{(0)}.$$

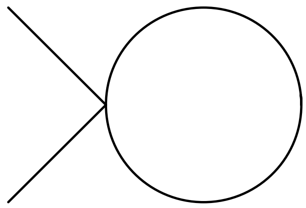


FIGURE 11. Self-energy graph in ϕ^4

Edges where two p_i, say p_1 and p_2, are paired together correspond to the simplest kind of graph. It splits off as a connected component and its contribution is of the form

$$(2\pi)^D\,\delta(p_1+p_2)\,(p_1^2+m^2)^{-1}. \tag{1.76}$$

Edges where one p_i is paired with one of the k_ℓ are *external.* Each p_i appears once and only once among the pairs, hence (1.69) holds. For each $j \in X$, each of the $k_i(j)$ appears exactly once in one of the pairs and (1.68) holds.
Notice that the two endpoints of an edge may well be the same, as in the tadpole graph of Figure 5. In the ϕ^3 theory this only occurs for tadpole graphs, but in other theories it occurs in more important graphs. For instance, in the case of the ϕ^4 theory this happens in the self-energy graph of Figure 11.

Feynman rules: Now that we know how to assign a graph to a pairing we can write down the Feynman rules which give the contribution of the pairing. For the moment, we ignore the combinatorial factor that accounts for the number of occurrences of the same graph from different pairings, which we describe later.
The Feynman rules assign:

1) To each external line $\ell \in \Gamma^{(1)}_{\rm ext}$ a propagator
$$\frac{1}{p_i^2+m^2}, \quad i \in \partial(\ell)$$

2) To each internal line $\ell \in \Gamma^{(1)}_{\rm int}$ a momentum variable $k = k_\ell$ and a propagator
$$\frac{1}{k^2+m^2}\,\frac{d^Dk}{(2\pi)^D}$$

3) To each vertex v with $\iota(v) = \frac{-z}{d!}\,\phi_E^d$ a momentum conservation
$$z\,(2\pi)^D\,\delta\Big(\sum_{\partial_0(\ell)=v} k_\ell - \sum_{\partial_1(\ell)=v} k_\ell\Big)$$

4) To each vertex v with $\iota(v) = \frac{-z}{2}\,(\partial\phi_E)^2$ a term
$$z\,(2\pi)^D\,k^2\,\delta\Big(\sum_{\partial_0(\ell)=v} k_\ell - \sum_{\partial_1(\ell)=v} k_\ell\Big).$$

In the last term the variable k is any of the k_ℓ involved in the momentum conservation (there are just two of them by (1.68)).

Unrenormalized values: The unrenormalized value $V(\Gamma)(p_1, \ldots, p_N)$ of the graph is the multiple integral obtained from the product of the four types of terms described above, assigned to Γ by the Feynman rules. It factorizes as a product

$$V(\Gamma) = \prod V(\Gamma_c) \tag{1.77}$$

over the connected components Γ_c of (the geometric realization of) Γ. In particular, it always contains an overall factor of the form

$$\epsilon(\Gamma) = \prod \epsilon(\Gamma_c), \tag{1.78}$$

where, for a connected graph, one lets

$$\epsilon(\Gamma) = (2\pi)^D \, \delta(\sum_j p_j) \prod_j \frac{1}{p_j^2 + m^2} \tag{1.79}$$

where p_j are the momentum variables assigned to the external lines and we assume for simplicity that all external lines are inward oriented. This can always be done, except when there is a single line, in which case one uses (1.76).

For a connected graph with N external legs, the number of free integration variables in the multiple integral $V(\Gamma)$ is

$$L(\Gamma) = \#\Gamma^{(1)}_{\rm int} - \#\Gamma^{(0)} + 1, \tag{1.80}$$

i.e. the number of internal lines minus the number of vertices plus one, which is the same thing as the first Betti number $b_1(|\Gamma|)$ of the geometric realization of Γ.

Symmetry factors: Let us determine the numerical factor in the contribution of (1.72) of the form

$$\frac{(-1)^n}{n!} \int \hat\phi(p_1) \cdots \hat\phi(p_N) \prod_{j \in X} \mathcal{M}_j \, d\mu.$$

The cardinality of X is n. We write each monomial in the form $\mathcal{M} = \dfrac{-z}{d!}\, \phi_E^d$, so that the minus signs cancel and we can ignore them. We also write the term $\prod_{j\in X} \mathcal{M}_j$ in the form

$$\prod_{j \in X} \mathcal{M}_j = \prod_I \mathcal{M}^{n_\iota},$$

where $\sum n_\iota = n$ occurs with a multiplicity $n!/\prod n_\iota!$ from the binomial expansion of $(\sum_I \mathcal{M})^n$. This gives an overall denominator $\delta = \prod n_\iota! \prod (d_\iota!)^{n_\iota}$, where d_ι is the order of the corresponding monomial $\mathcal{M}$.

One sees that δ is the order of the group Δ of permutations of the terms $\hat\phi(k_i(j))$ in (1.75) that respect the map $X \to I$. In particular, this group Δ acts on the pairings and the orbit $\Delta(\pi)$ of a given pairing gives all pairings of the same type.

Thus, we see that the symmetry factor is given by the cardinality of the isotropy group

$$\sigma(\Gamma) = \#\{g \in \Delta \,|\, g(\pi) = \pi\}. \tag{1.81}$$

Since the construction of the graph Γ from the pairing π is natural, the isotropy group is the group of automorphisms of the graph, so that

$$\sigma(\Gamma) = \#\mathrm{Aut}(\Gamma). \tag{1.82}$$

It accounts for repetitions, as usual in combinatorics. Notice that we do not take into account the *orientation* of the edges, i.e. we do not consider two graphs which only differ by a choice of orientations of their edges as distinct.

3.4. Connected Green's functions.

There are standard procedures that allow one to simplify the combinatorics of the Feynman graphs. First, one can combine the Green's functions (1.25) in a single generating function. This takes the form of a Fourier transform

$$
\begin{aligned}
Z(J) = & \mathcal{N} \int \exp\left(i \tfrac{S(\phi)+\langle J,\phi\rangle}{\hbar}\right) \mathcal{D}[\phi] \\
= & \textstyle\sum_{N=0}^{\infty} \frac{i^N}{N!} \int J(x_1)\cdots J(x_N)\, G_N(x_1,\ldots,x_N)\, dx_1\cdots dx_N,
\end{aligned} \tag{1.83}
$$

where the *source* J is an element of the linear space dual to that of the classical fields ϕ, with $\langle J,\phi\rangle = \int J(x)\phi(x)dx$.

The first simplification of the class of graphs that need to be considered is obtained by passing to the logarithm of $Z(J)$,

$$
\begin{aligned}
iW(J) &= \log(Z(J)) \\
&= \sum_{N=0}^{\infty} \frac{i^N}{N!} \int J(x_1)\cdots J(x_N) G_{N,c}(x_1,\ldots,x_N) dx_1\cdots dx_N.
\end{aligned} \tag{1.84}
$$

This is the generating function for the *connected* Green's functions $G_{N,c}$. This means that, at the formal combinatorial level, the perturbative expansion

$$
\begin{gathered}
G_N(x_1,\ldots,x_N) = \\
\sum_{\Gamma} \int \frac{V(\Gamma)(p_1,\ldots,p_N)}{\sigma(\Gamma)}\, e^{i(x_1\cdot p_1+\cdots+x_N\cdot p_N)} \prod_j \frac{dp_j}{(2\pi)^D}
\end{gathered} \tag{1.85}
$$

of the original Green's functions (i.e. the Minkowski space analogue of (1.65)) is over all graphs (including the non-connected ones). By taking the log and passing to the expression $W(J)$ in (1.84), one simply drops from the counting all the disconnected graphs, so that

$$
\begin{gathered}
G_{N,c}(x_1,\ldots,x_N) = \\
\sum_{\Gamma \text{ connected}} \int \frac{V(\Gamma)(p_1,\ldots,p_N)}{\sigma(\Gamma)}\, e^{i(x_1\cdot p_1+\cdots+x_N\cdot p_N)} \prod_j \frac{dp_j}{(2\pi)^D}.
\end{gathered} \tag{1.86}
$$

At the same time, dividing by the normalization factor $\mathcal{N}$ in (1.83) eliminates all the "vacuum bubbles", that is, all the graphs that do not have any external legs.

The number L of loops in a connected graph determines the power $\hbar^{L-1}$ of the unit of action that multiplies the corresponding term, so that (1.84) has the form of a semiclassical expansion. Indeed, if we keep track of the powers of $\hbar$ in the perturbative expansion, we see that each propagator contributes one factor of $\hbar$ (since it is the inverse of the free part of the action) and each interaction term a factor of $\hbar^{-1}$. Thus, each connected graph Γ with N external legs inherits a factor of $\hbar^{N+L(\Gamma)-1}$. The source J appears in the action with a coefficient of $\hbar^{-1}$, hence we are left with a factor of $\hbar^{L(\Gamma)-1}$ for the contribution of each graph to $W(J)$.

In a similar manner, the generating function of Euclidean Green's functions is given by

$$
\begin{aligned}
&Z(J_E) = \\
&\mathcal{N} \int \exp\left(-\tfrac{S(\phi_E)-\langle J_E,\phi_E\rangle}{\hbar}\right) \mathcal{D}[\phi_E] = \\
&\sum_{N=0}^{\infty} \frac{1}{N!} \int J_E(x_1)\cdots J_E(x_N)\, S_N(x_1,\ldots,x_N)\, dx_1\cdots dx_N,
\end{aligned} \tag{1.87}
$$

where the *source* J_E is an element of the linear space dual to that of the Euclidean classical fields ϕ_E. The generating function for the connected Euclidean Green's functions is of the form

$$
\begin{aligned}
&W(J_E) = \log(Z(J_E)) = \\
&\sum_{N=0}^{\infty} \frac{1}{N!} \int J_E(x_1)\cdots J_E(x_N) S_{N,c}(x_1,\ldots,x_N)\, dx_1\cdots dx_N\,,
\end{aligned} \tag{1.88}
$$

and is similar to the free energy of a statistical mechanical system in the presence of the source J_E.

Suppose given a formal expansion in terms of connected graphs and symmetry factors

$$
W = \sum_{\Gamma\,\text{connected}} \frac{V(\Gamma)}{\sigma(\Gamma)}. \tag{1.89}
$$

The relations that give the equality

$$
e^W = \sum_{\Gamma} \frac{V(\Gamma)}{\sigma(\Gamma)}, \tag{1.90}
$$

with the sum over all (not necessarily connected) graphs, are of the form

$$
V(\Gamma_1 \cup \Gamma_2) = V(\Gamma_1)\, V(\Gamma_2)\,, \quad \text{with} \quad \sigma(\Gamma) = \prod n_j! \prod \sigma(\Gamma_j)^{n_j} \tag{1.91}
$$

where the Γ_j are the distinct graphs that appear as components of Γ and the n_j are the multiplicities of their occurrences.

Using this general principle one obtains

$$
W(J_E) = \log Z(J_E) = \sum_{\Gamma \text{ connected}} \frac{V(\Gamma)(J_E)}{\sigma(\Gamma)}, \tag{1.92}
$$

where

$$
V(\Gamma)(J) = \frac{1}{N!} \int \hat{J}(p_1)\cdots \hat{J}(p_N)\, V(\Gamma(p_1,\ldots,p_N)) \prod_j \frac{dp_j}{(2\pi)^D}\,. \tag{1.93}
$$

3.5. The effective action and one-particle irreducible graphs.

There is then a second key simplification of the combinatorics of graphs which makes it possible to reduce the computations to connected graphs which are *one-particle irreducible* (1PI).

Definition 1.3. *Let Γ be a connected Feynman graph of a given QFT. Then Γ is a* one-particle irreducible *(1PI) graph if the following holds.*

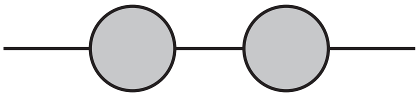


FIGURE 12. One-particle reducible graph

(1) Γ *is not a tree.*
(2) Γ *cannot be disconnected by cutting a single edge.*

Since in the following we work mostly in the Euclidean case, we drop the distinction between ϕ_E and ϕ and we simply write ϕ for the scalar field in the Euclidean theory.

The second simplification we now introduce is achieved by passing to the *effective action* $S_{\text{eff}}(\phi)$. We describe it in the Euclidean context. There are similar formulas (with suitable powers of i) in the Minkowski signature. By definition, $S_{\text{eff}}(\phi_E)$ is the Legendre transform of $W(J_E)$, namely it is defined by

$$S_{\text{eff}}(\phi) = \langle \phi, J_E \rangle - W(J_E)|_{J_E = J(\phi)}, \tag{1.94}$$

where J_E is chosen so that the functional (variational) derivative $\delta W/\delta J$ of W with respect to J at J_E is equal to ϕ (i.e. so that the right-hand side of (1.94) is stationary). Such a J_E can be uniquely determined in perturbation theory. In the case of free fields, with $S(\phi) = S_0(\phi)$, the effective action is the same as the original action, with the possible addition of an irrelevant constant. In the general case, where an interaction term $S_{\text{int}}(\phi)$ is present in the action, the effective action $S_{\text{eff}}(\phi)$ is a non-linear functional of the classical fields, which constitutes the basic unknown in a given quantum field theory.

The key result on the effective action is that it can be computed by dropping all graphs that can be disconnected by the removal of one edge, as in the case illustrated in Figure 12, where the shaded areas are a shorthand notation for an arbitrary graph with the specified external leg structure. One can reduce in this way the combinatorics of graphs and restrict attention to 1PI graphs only. The individual contribution of 1PI graphs to the non-linear functional $S_{\text{eff}}(\phi)$ can be spelled out very concretely as follows. Up to a sign and a symmetry factor (Theorem 1.5) the contribution to the effective action of the 1PI graph Γ is of the form (cf. (1.93))

$$U(\Gamma)(\phi) = \frac{1}{N!} \int_{\sum p_j = 0} \hat{\phi}(p_1) \cdots \hat{\phi}(p_N)\, U(\Gamma(p_1, \ldots, p_N)) \prod_j \frac{dp_j}{(2\pi)^D}. \tag{1.95}$$

Here N is the number of external legs of the graph Γ, while $\hat{\phi}$ denotes the Fourier transform of ϕ (we are integrating in momentum space). Formula (1.95) should be understood as the pairing of the distribution $U(\Gamma(p_1, \ldots, p_N))$ with the Fourier transforms $\hat{\phi}(p_i)$. The distribution $U(\Gamma(p_1, \ldots, p_N))$ is in fact given by a smooth function of the external momenta p_i and is specified by the *Feynman rules* of the theory. The $U(\Gamma(p_1, \ldots, p_N))$ is referred to as the *unrenormalized value* of the graph Γ with assigned external momenta $(p_1, \ldots, p_N)$.

FIGURE 13. Connected graph with two external legs

DEFINITION 1.4. *Let* Γ *be a 1PI graph. Its unrenormalized value, denoted by* $U(\Gamma(p_1,\ldots,p_N))$*, is given by* $V(\Gamma)$ *with external propagators and* δ*-function removed, so that*

$$V(\Gamma(p_1,\ldots,p_N)) = \epsilon(\Gamma)\, U(\Gamma(p_1,\ldots,p_N)). \tag{1.96}$$

It is only defined for $\sum_j p_j = 0$ and it is of the general form

$$U(\Gamma(p_1,\ldots,p_N)) = \int I_\Gamma(k_1,\ldots k_L, p_1,\ldots p_N)\, d^D k_1 \cdots d^D k_L, \tag{1.97}$$

where $L = b_1(|\Gamma|)$ is the loop number of the graph (the first Betti number) and the $k_i \in \mathbb{R}^D$, with $i = 1,\ldots,L$, are momentum variables assigned to the internal edges.

THEOREM 1.5. *The effective action is given by the formal series*

$$S_{\text{eff}}(\phi) = S(\phi) - \sum_{\Gamma\in\,\text{1PI}} \frac{U(\Gamma)(\phi)}{\sigma(\Gamma)}\,. \tag{1.98}$$

PROOF. If we keep explicit count of the powers of $\hbar$, then (1.98) becomes

$$\frac{S_{\text{eff}}(\phi)}{\hbar} = \frac{S(\phi)}{\hbar} - \frac{1}{\hbar}\sum_{\Gamma\in\,\text{1PI}} \hbar^L \frac{\Gamma(\phi)}{\sigma(\Gamma)},$$

where the loop number $L(\Gamma)$ of 1PI graphs is ≥ 1 since we excluded tree graphs. Thus, the effective action $S_{\text{eff}}(\phi)$ encodes the quantum corrections to the original action. Its knowledge makes it possible to calculate Green's functions in a direct manner. Namely, one obtains the same answer by computing at *tree level* with the effective action $S_{\text{eff}}(\phi)$ as by computing the full perturbative expansion with the original action $S(\phi)$,

$$\mathcal{N}\int P(\phi)\, e^{-S(\phi)}\, \mathcal{D}[\phi] = \mathcal{N}\int_{\mathbf{tree-level}} P(\phi)\, e^{-S_{\text{eff}}(\phi)}\, \mathcal{D}[\phi]\,. \tag{1.99}$$

To compute the right-hand side one isolates the quadratic part of the functional $S_{\text{eff}}(\phi)$ and then applies perturbation theory, but one drops all non-tree graphs. The basic facts that show how this works are the following.

a) Let Γ be a connected graph with two external legs. Then one can write Γ uniquely as a tree graph with interaction vertices given by 1PI graphs Π with two external legs.
b) The sum of the contributions of all the connected graphs as above is given by

$$\sum_{\Gamma} \frac{V(\Gamma(p,-p))}{\sigma(\Gamma)} = \frac{1}{p^2 + m^2 - \sum_\Pi \frac{U(\Pi(p,-p))}{\sigma(\Pi)}}\,. \tag{1.100}$$

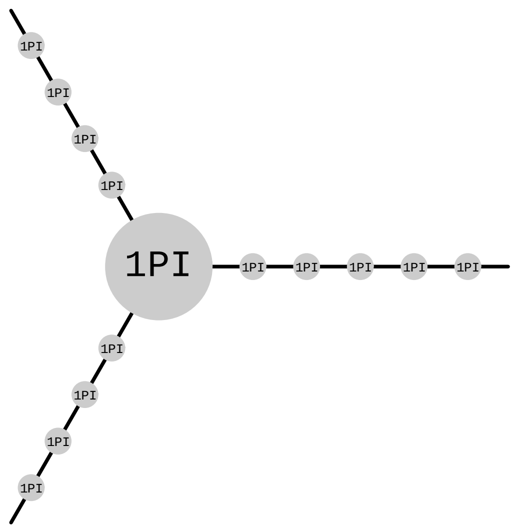


FIGURE 14. Connected graph with three external legs

The reason to drop the external momenta in the definition 1.4 of $U(\Pi)$ can be understood very simply from the evaluation of $V(\Gamma(p,-p))$ for a graph of the form $\Pi\star\Pi\star\cdots\star\Pi$ as in Figure 13, where $\star$ here denotes the concatenation product. Such a graph gives, up to symmetry factors, the expression

$$V(\Gamma(p,-p)) = (p^2+m^2)^{-q-1}\, U(\Pi(p,-p))^q.$$

Summing the geometric series (allowing for different choices of Π) gives the required equality b). Extending a) to connected graphs with an arbitrary number of external legs as shown in the case of graphs with three external legs in Figure 14 proves (1.99).

This suffices to prove Theorem 1.5. Indeed, computing the functional integral in the right-hand side of (1.99) at tree-level, means that one just applies the stationary phase method to the functional integral with the amplitude

$$\exp\left(-\frac{S_{\text{eff}}(\phi)}{T}\right), \tag{1.101}$$

when the parameter T (which has nothing to do with $\hbar$) tends to 0. We know from (1.99) that this yields the same result as the full calculation of the integrals with respect to the original Euclidean action (1.45).
Introducing a source term J implies that $W(J)$ is the Legendre transform of $S_{\text{eff}}(\phi)$ and, by reciprocity of the Legendre transform, that Theorem 1.5 holds. $\square$

The computation of the non-linear functional of classical fields given by the effective action is the basic first problem in QFT. If one knows the effective action, one can obtain from it all the Green's functions by tree level calculations, applying the stationary phase approximation. This makes it possible to compute the physical observables such as the S-matrix elements.

3.6. Physically observable parameters.

The main idea of the renormalization process is the distinction between the bare parameters that enter in the Lagrangian and the physically observable parameters. As a prerequisite to understanding renormalization one needs to define the physically observable parameters. This is done in terms of the effective action in a way that makes it possible to compute them in a direct manner. We explain briefly how.

Physical mass: The notions of particle and of mass in QFT are understood through the classification of irreducible representations of the Poincaré group. In the case at hand of a scalar field ϕ, the spin vanishes and the relevant irreducible representations are induced from the mass-shell orbits of the Lorentz group acting on momentum space. For each $\mu > 0$ one lets $\mathcal{H}_\mu$ be the Hilbert space of square-integrable functions on the mass shell $k^2 = \mu^2, k_0 > 0$, and defines an irreducible representation of the Poincaré group. In the free quantum field theory the particle description is in terms of the Fock space

$$\mathcal{F}_m = \bigoplus_{n=0}^{\infty} S^n \mathcal{H}_m, \tag{1.102}$$

where $S^n\mathcal{H}$ is the n-th symmetric tensor power of the one-particle Hilbert space $\mathcal{H}_m$. The natural action of the Poincaré group on the Fock space is no longer irreducible and decomposes as an integral of irreducible ones with a discrete part where the mass m appears as the smallest eigenvalue with a mass gap. Still in the case of the free field, one can compute the mass m by looking at the connected Minkowski space Green's functions. Since there is no interaction, the only connected graph is the trivial two-point graph and the only non-zero connected Green's function is the two-point function given by (1.42)

$$\begin{aligned} G_{2,c}(x_1, x_2) &= \int \frac{i}{p^2 - m^2 + i\epsilon}\, e^{i(x_1 - x_2)\cdot p}\, \frac{d^D p}{(2\pi)^D} \\ &= \Delta_+(x_1 - x_2, m^2). \end{aligned} \tag{1.103}$$

In the general interacting case one has the Källen–Lehmann spectral decomposition which asserts the existence of a positive density $\sigma(u)$ for $u > m^2_{\text{phys}}$ such that:

$$\begin{aligned} Z^{-1}\, G_{2,c}(x_1, x_2) = &\ \Delta_+(x_1 - x_2, m^2_{\text{phys}}) \\ &+ \int_{m^2_{\text{phys}}}^{\infty} \Delta_+(x_1 - x_2, u)\, \sigma(u)\, du \end{aligned} \tag{1.104}$$

where the normalization factor Z is described in more detail in (1.109) below. The formula (1.104) shows that the physical mass m_{phys} can be read off from the pole structure of the Fourier transform of the connected Minkowski space Green's function $G_{2,c}$. It has an isolated simple pole at $p^2 = m^2_{\text{phys}}$, which signals the presence of a particle of mass m_{phys} and makes it possible to apply scattering theory.
Now by Theorem (1.5) we know that $G_{2,c}$ is given by the tree-level calculation using the effective action. In fact, (1.100) gives for the (Fourier transform of) the Euclidean connected two-point function the formula

$$S_{2,c}(p) = \frac{1}{p^2 + m^2 - \sum_{\Pi} U(\Pi(p, -p))/\sigma(\Pi)}. \tag{1.105}$$

This shows that the physical mass m_{phys} can be computed in perturbation theory by solving the equation

$$p^2 + m^2 - \sum_{\Pi} \frac{U(\Pi(p,-p))}{\sigma(\Pi)} = 0 \quad \text{for} \quad p^2 = -m_{\text{phys}}^2. \tag{1.106}$$

The expression

$$\Pi(p^2) = \sum \frac{U(\Pi(p,-p))}{\sigma(\Pi)} \tag{1.107}$$

is called the *self-energy*.

Physical field and scattering matrix: The presence of the normalization factor Z in the Källen–Lehmann formula (1.104) can be eliminated by a rescaling of the quantum field ϕ. One simply defines

$$\phi' = Z^{-1/2}\, \phi. \tag{1.108}$$

The factor Z is the residue

$$Z = \operatorname{Res}_{p^2 = -\, m_{\text{phys}}^2} S_{2,c}(p) \tag{1.109}$$

of (1.105)) at the pole $p^2 = -\, m_{\text{phys}}^2$.

The basic formula which makes contact with experiments is the scattering formula of Lehmann–Symanzik–Zimmermann (LSZ), which gives the scattering matrix S in the form

$$\begin{gathered} \langle k_1, \dots, k_s | S - 1 | k_{s+1}, \dots, k_r \rangle = \\ i^{-r} Z^{-r/2} \textstyle\prod_j (k_j^2 - \, m_{\text{phys}}^2)\, G_N(-k_1, \dots, -k_s, k_{s+1}, \dots, k_r). \end{gathered} \tag{1.110}$$

This says that a matrix element in $S - 1$ is a momentum space Green's function (here in Minkowski space), with external propagators removed, rescaled by a suitable power of the residue of the propagator at the square of the physical mass.

Effective coupling constant: The knowledge of the scattering matrix makes it possible to determine the strength of the coupling, by looking at the matrix elements involving $r = 3$ particles. Consider the ϕ^3 theory with Euclidean Lagrangian

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 + \frac{m^2}{2}\phi^2 - \frac{g}{6}\phi^3, \tag{1.111}$$

where we change the sign of the coupling constant to eliminate the minus signs in the perturbative expansion. One defines the effective coupling constant from the cubic term in the effective action as in the LSZ formula (1.110), using the overall factor of $Z^{-r/2}$. This gives

$$g_{\text{eff}} = Z^{-3/2} \left(g + \sum \frac{U(\Gamma(0,0,0))}{\sigma(\Gamma)} \right), \tag{1.112}$$

where the sum is over all 1PI graphs with three external lines, each with a zero momentum.

3.7. The physics idea of renormalization.

The rule is that one computes physical quantities in terms of the observable parameters such as the physical mass, not the bare parameters such as the bare mass. The main point is not that of getting finite quantities (any regularization procedure does that), but to eliminate the regularization parameters (such as the parameter Λ in the cutoff) by a suitable addition of *counterterms* to the original Lagrangian. The main requirements on the counterterms are the following:

- Locality: they are polynomials in the fields and their derivatives.
- Finiteness: The number of monomials in the Lagrangian remains finite after the introduction of all the necessary counterterms.

The reason for the finiteness requirement is that each new counterterm leaves behind a free parameter of the theory (for instance the physical mass). There is no way such a parameter can be fixed other than by its experimental value, hence one gets not a single theory but a family of theories parameterized by as many free parameters as the counterterms added to the Lagrangian. Since a theory with an infinite number of free parameters has virtually no predictive power (it can always accommodate any new experimental result) the finiteness of the number of counterterms is required.

We now explain more carefully the meaning of the three types of counterterms needed to renormalize the theory in the above generic example, namely

- Mass counterterm (renormalization)
- Field strength counterterm (renormalization)
- Coupling constant counterterm (renormalization)

Mass renormalization: Let us consider the simple example of the self-energy graph (1.57) for $D = 4$. We write its contribution to the self-energy (1.107). It is given by

$$\Pi(p^2) = \frac{1}{2}g^2\,(2\pi)^{-4}\int \frac{1}{k^2+m^2}\,\frac{1}{((p+k)^2+m^2)}\,d^4k. \tag{1.113}$$

In physics, the idea of how to handle this divergent expression is to distinguish between physical and non-physical mass parameters. On the one hand we have the bare mass parameter

$$m_0 = \text{ bare mass},$$

which specifies the mass term in the Lagrangian. It is the parameter that appears in the term $\frac{1}{2}m_0^2\phi^2$ in the Euclidean Lagrangian. On the other hand, one has the measured experimental mass m_{phys}, which arises as the position of the pole in the full propagator, i.e. as the zero of its inverse

$$p^2 + m^2 - \Pi(p^2) = 0 \quad \text{for } \; p^2 = -\,m_{\text{phys}}^2.$$

Let $m = m_{\text{phys}}$ denote the physical mass. After introducing a cutoff parameter Λ one writes

$$m_0^2 = m^2 - \delta m^2(\Lambda).$$

In this way one views the term $-\delta m^2(\Lambda)\frac{\phi^2}{2}$ as an interaction term in the Lagrangian. Thus, one does perturbation theory with the free part given by the (Euclidean) propagator $(p^2+m^2)^{-1}$ and one determines δm^2 in perturbation theory as a function of a regularization parameter, such as a cutoff Λ, in such a way that it cancels the divergence but does not spoil the location of the pole, so that the self-energy Π satisfies the *renormalization condition*

$$\Pi(-m^2) = 0. \tag{1.114}$$

The interaction term $-\delta m^2 \frac{\phi^2}{2}$ contributes by the tree graph of Figure 10 to the self-energy. This contributes an additional term of δm^2 to the self-energy. Thus, at this order it is easy to solve (1.114). Namely, one takes

$$\delta m^2 = -\Pi(-m^2).$$

Thus, at order g^2 in the perturbation theory, the renormalized self-energy is of the form

$$\Pi_{\mathrm{ren}}(p^2) = \Pi(p^2) - \Pi(-m^2). \tag{1.115}$$

This is now a *finite* quantity, which can be computed using any regularization method such as a cutoff, giving the answer

$$\Pi_{\mathrm{ren}}(p^2) = \frac{g^2}{32\pi^2} \int_0^1 \log\left(\frac{x(1-x)p^2 + m^2}{m^2(1+x-x^2)}\right) dx. \tag{1.116}$$

Besides the mass counterterm, there are two other types of counterterms, allowing for the renormalization of the field strength and the coupling constants.

Field strength renormalization: We saw in the Lehmann–Symanzik–Zimmermann scattering formula (1.110) that the residue of the two-point connected Green's function at the pole enters explicitly in the expression for the scattering matrix, so that the physical field ϕ' is related to the bare field ϕ by equation (1.108). The change of variable formula $\phi \to Z^{1/2}\,\phi'$ in the Euclidean functional integral

$$\begin{gathered} \mathcal{N} \int \exp(-S(\phi))\, \phi(x_1) \cdots \phi(x_N) \mathcal{D}[\phi] = \\ Z^{N/2} \mathcal{N}' \int \exp(-S(Z^{1/2}\,\phi'))\, \phi'(x_1) \cdots \phi'(x_N) \mathcal{D}[\phi'] \end{gathered} \tag{1.117}$$

shows that the distinction between the physical field and the bare field makes it possible to modify the kinetic term in the Lagrangian. When expressed in terms of ϕ', this takes the form

$$\mathcal{L}' = \frac{1}{2} Z (\partial\phi')^2 + \cdots$$

Writing $Z = 1 - \delta Z(\Lambda)$, one then chooses the *field strength* counterterm $\delta Z(\Lambda)$ in such a way that the residue at the pole remains equal to 1. In terms of the self-energy this means that the derivative vanishes at $-m^2$, i.e. that

$$\partial\Pi(-m^2) = 0. \tag{1.118}$$

In the example above only a finite field strength renormalization is required to compensate for the contribution of (1.116) to the self-energy. Notice that, since (1.115) is a shift by an additive constant, the finiteness of the field strength renormalization in this example comes from the absence of a momentum dependence in the divergence of the integral (1.113). If we look at the same ϕ^3 theory in dimension $D = 6$, then the divergence of the integral (1.113) is not canceled by the mass counterterm alone. However, one can replace (1.115) by

$$\Pi_{\mathrm{ren}}(p^2) = \Pi(p^2) - \Pi(-m^2) + \Pi'(-m^2)\,(p^2 - m^2), \tag{1.119}$$

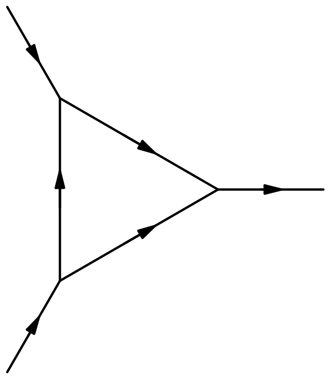


FIGURE 15. Graph with three external legs

where Π' is the derivative of Π with respect to p^2. Exactly as for mass renormalization, one then computes the counterterm $\delta Z(\Lambda)$ using the renormalization condition (1.118).

Coupling constant renormalization: In the same ϕ^3 theory in dimension $D = 6$ the graph of Figure 15 is 1PI and gives a contribution of the form

$$g^3\,(2\pi)^{-6} \int \frac{1}{k^2+m^2}\,\frac{1}{(p+k)^2+m^2}\,\frac{1}{(q-k)^2+m^2}\,d^6k. \tag{1.120}$$

The distinction between the bare coupling constant and the effective coupling constant of (1.112) gives us the freedom to write the bare coupling in the form $g + \delta g$ and to introduce the counterterm δg. The last renormalization prescription is that the effective coupling (1.112) is fixed by its experimental value. The full (Euclidean) Lagrangian with the counterterms now takes the form

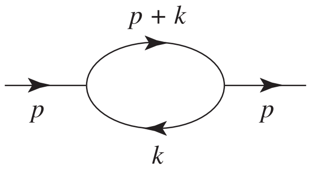
$$\mathcal{L}_E = \frac{1}{2}(\partial\phi)^2(1-\delta Z) + \left(\frac{m^2-\delta m^2}{2}\right)\phi^2 - \frac{g+\delta g}{6}\phi^3. \tag{1.121}$$

The renormalization scheme described above is called the mass-shell scheme and is the scheme which is easiest to interpret in physics terms. One computes physically observable quantities such as the S-matrix elements not in terms of the bare parameters but in terms of the physical parameters, which in the example are the mass, field strength and coupling constant. This is done by computing the bare parameters as formal power series in the coupling so that the renormalization conditions such as (1.114) and (1.118) are fulfilled while the effective coupling constant is held fixed at its observed value.

4. Dimensional regularization

As we saw above, the cutoff method and the mass-shell scheme allow for a natural explanation of the basic physics content of renormalization. However, in general one needs more refined *regularization* procedures for the integrals involved in the expression of the $U(\Gamma(p_1,\ldots,p_N))$, which do not break Lorentz invariance. Moreover, *renormalization schemes* which are compatible with the massless case are needed in general. We use the *regularization scheme* known as *Dimensional regularization* (DimReg), for the reason that, when combined with *Minimal Subtraction* (MS), it is the most widely used renormalization scheme in explicit perturbative calculations in quantum field theory and elementary particle physics. The method was developed by 't Hooft and Veltman [**169**], with the aim of treating perturbative calculations involving nonabelian gauge theories, theories with fermions, and anomalies. As we explain in Section 19 of this Chapter, DimReg also turns out to be very closely related to noncommutative geometry.

Before giving a general definition, we explain what it means to apply DimReg in the simple example of the self-energy graph (1.57). Its contribution as a 1PI graph is given (up to $g^2\,(2\pi)^{-D}$) by



$$(1.122) \qquad \to \int \frac{1}{k^2+m^2}\,\frac{1}{((p+k)^2+m^2)}\,d^Dk.$$

In fact, recall that for 1PI graphs one does not have a contribution from the external propagators (cf. (1.96)).

One first introduces the *Schwinger parameters.* These replace each propagator of the form $\frac{1}{P}$ by $\int_0^\infty e^{-sP}ds$, introducing in this way an additional parameter s that will then be eliminated by integration. This means, in the case of (1.122), that we use the simple identity

$$(1.123) \qquad \frac{1}{k^2+m^2}\,\frac{1}{(p+k)^2+m^2} = \int_{s>0,\ t>0} e^{-s(k^2+m^2)-t((p+k)^2+m^2)}\,ds\,dt.$$

One diagonalizes the quadratic form $-Q(k)$ in the exponential, which is then written as

$$-Q(k) = -\lambda\,((k+xp)^2 + ((x-x^2)p^2+m^2)), \qquad \text{with} \;\; s=(1-x)\lambda \;\text{ and }\; t = x\,\lambda.$$

Thus, after changing the order of integration, one obtains a Gaussian integral in dimension D in the variable $q = k + xp$, of the form

$$(1.124) \qquad \int e^{-\lambda\,q^2}\,d^Dq \;=\; \pi^{D/2}\,\lambda^{-D/2}\,.$$

In this way we have rewritten the unrenormalized value $U(\Gamma)$ of the graph (1.122) in dimension D in the form

$$(1.125) \qquad \int_0^1\int_0^\infty e^{-(\lambda(x-x^2)p^2+\lambda\,m^2)}\int e^{-\lambda\,q^2}\,d^Dq\;\lambda\,d\lambda\;dx$$

$$= \pi^{D/2}\int_0^1\int_0^\infty e^{-(\lambda(x-x^2)p^2+\lambda\,m^2)}\,\lambda^{-D/2}\,\lambda\,d\lambda\,dx$$

$$(1.126) \qquad = \pi^{D/2}\,\Gamma(2-D/2)\int_0^1 ((x-x^2)p^2+\,m^2)^{D/2-2}\,dx.$$

The remaining integral can be computed in terms of hypergeometric functions. We do not need to write it out explicitly, since what matters here is the fact that the divergence of the integral (1.122) is now attributed to the presence of a pole of the

Gamma function $\Gamma(2-D/2)$ at a point $D \in 4+2\mathbb{N}$. Moreover, all the terms that appear in the expression (1.126) now continue to make sense when the integer D is replaced by a complex number $D-z \in \mathbb{C}$, with z in a small disk $\Delta \subset \mathbb{C}$ centered at zero. Notice also that, in (1.126), the coefficient of the pole of $\Gamma(2-D/2)$ at $D \in 4+2\mathbb{N}$ is a polynomial in p, hence its Fourier transform is a *local* term.

This example exhibits all the characteristics of the DimReg method. In general, the main idea is to use the basic formula (1.124) to *define* the meaning of the integral when D is no longer an integer.

More precisely, according to the Feynman rules, the expression I_Γ of (1.97) is typically of the form

$$I_\Gamma(\kappa) = \frac{P(\kappa,p)}{\prod_j F_j(\kappa,p)^2 + m_j^2}, \tag{1.127}$$

with $\kappa = (k_1,\ldots,k_L)$ and $p = (p_1,\ldots,p_N)$. Here P is a polynomial, and the F_j, $j = 1,\ldots,L+N$ are linear forms on $\mathbb{R}^{L+N}$. It is better to be more specific. In all cases of interest the $F_j(\kappa,p)$ turn out to be of the form

$$F_j(\kappa,p) = \sum_i a_{j,i}\,(\kappa_j)_i + \sum_\ell b_{j,\ell}\,(p_j)_\ell, \tag{1.128}$$

where the coefficients are integers. We use the same formula when the $\kappa_j = (\kappa_j)_i$ and the $p_j = (p_j)_\ell$ are no longer scalars but elements of a fixed Euclidean space E of dimension D. We use the notation v^2 for the square of the Euclidean norm of an element $v \in E$. Thus $F_j(\kappa,p)^2$ is now a quadratic form on E^{L+N} and can be written using an orthogonal basis e_r of E as the sum of the D expressions

$$F_j(\kappa,p)^2 = \sum_r (e_r \circ F_j(\kappa,p))^2, \tag{1.129}$$

where e_r is viewed as a linear form on E so that each $e_r \circ F_j(\kappa,p)$ is a linear form on E^{L+N}.

To compute the integral associated to (1.127) one uses the Schwinger parameters as before, rewriting the integral (1.97) in the form

$$\int P(\kappa,p)\left(\int e^{-\sum t_j(F_j(\kappa,p)^2+m_j^2)}\,dt\right) d^D\kappa,$$

with $d^D\kappa = d^Dk_1\cdots d^Dk_L$ and $dt = \prod_j dt_j$, integrated over $t_j > 0$. Formally, one permutes the order of integration and sets the above equal to

$$\int \left(e^{-\sum t_j m_j^2}\int P(\kappa,p)e^{-\sum t_j F_j(\kappa,p)^2}d^D\kappa\right) dt.$$

One then treats the integral in $d^D\kappa$ as a Gaussian integral. One writes

$$\sum t_j F_j(\kappa,p)^2 = A_E(t)(q) + \sum t_j F_j(\kappa(t),p)^2\,, \quad \text{with } \; q = \kappa - \kappa(t),$$

where the quadratic form $A_E(t)$ on E^L is given by

$$A_E(t)(q) = \sum t_j F_j(q,0)^2 \tag{1.130}$$

and the minimum $\kappa(t)$ of $\sum t_j F_j(\kappa,p)^2$ depends rationally on t.

In fact, by construction $\kappa(t)$ is of the form $A(t)^{-1}\ell(t,p)$, where $\ell(t,p)$ is a bilinear function of t and p. After a change of variables $\kappa \mapsto q = \kappa - \kappa(t)$ one is left with the integral of a polynomial in q with respect to the Gaussian $e^{-A_E(t)(q)}\, d^D q$.
Using integration by parts (1.38), the only integration in $d^D q$ to be performed is that of the Gaussian, which gives

$$\int e^{-A_E(t)(q)}\, d^D q = \pi^{LD/2}\, \det(A(t))^{-D/2},$$

where $A(t)$ is the quadratic form (1.130) but in the *one-dimensional case*. This means that one uses formula (1.130) to define the quadratic form $A(t)$ on $\mathbb{R}^L$. One checks that one gets the correct result using (1.129).
Both the integration by parts and the dependence $\kappa(t) = A(t)^{-1}\ell(t)$ introduce denominators, given by powers of $\det(A(t))$. Thus, the end result is an integral of the form

$$\pi^{LD/2} \int e^{-\sum t_j m_j^2 - \sum t_j F_j(\kappa(t),p)^2}\, f(p,t)\, \det(A(t))^{-D/2-n}\, dt\,, \tag{1.131}$$

where f depends polynomially on t and p.
The integral obtained in this way is divergent, because of the singularity coming from the zeros of the polynomial $Q(t) = \det(A(t))$, but all the terms that appear in (1.131) now make sense when D is continued to $D - z \in \mathbb{C}$.

Thus, we have formally obtained an integral in dimension $D - z$ which replaces the integral (1.97). Namely, we have

$$\begin{aligned} &U^z(\Gamma(p_1,\ldots,p_N)) = \\ &\int I_\Gamma(k_1,\ldots k_L, p_1,\ldots p_N) d^{D-z}k_1 \cdots d^{D-z}k_L := \\ &\int e^{-\sum t_j m_j^2 - \sum t_j F_j(\kappa(t),p)^2} f(p,t)\, \det(A(t))^{-(D-z)/2-n}\, dt, \end{aligned} \tag{1.132}$$

with the last line obtained as we just described. The integral is convergent when the real part of z is sufficiently large.

We now discuss the existence of a meromorphic continuation to the complex plane. The mathematical treatment, cf. e.g. [**128**] is based on the following result of Bernstein [**17**], applied to the polynomial $\det(A(t))$.

LEMMA 1.6. *Let $Q(t)$ be a polynomial in n variables. There exists a polynomial $q(D)$ and a polynomial differential operator $L(D)$ in n variables, whose coefficients are polynomials in D, such that*

$$L(D)\, Q^{-D/2} = q(D)\, Q^{-D/2-1}\,, \quad \forall D.$$

PROOF. We first give a proof in the "generic" case, i.e. assuming that the hypersurface $Q = 0$ is smooth. In that case, by Hilbert's Nullstellensatz, the ideal generated by the polynomial Q and its first derivatives $\partial_j Q$ contains the constant polynomial 1. Thus, one can find polynomials A and A_j such that

$$A\, Q + \sum_j A_j\, \partial_j Q = 1.$$

One just lets

$$L(D) = -\frac{D}{2}\,A + \sum_j A_j\,\partial_j \quad \text{and} \quad q(D) = -\frac{D}{2}$$

In the general case ([**17**]) one considers the space $\mathcal{P}$ of expressions of the form $p(t,D)\,Q^{-D/2-k}$ where $p(t,D)$ is a polynomial in t whose coefficients are rational functions of D. It is a $\mathcal{D}$-module, i.e. a module over the ring of differential operators $P(t,\partial,D)$ whose coefficients are polynomials in t and rational functions of D. The main point is to show that, as a $\mathcal{D}$-module, $\mathcal{P}$ is finitely generated. It follows that the increasing filtration $\mathcal{F}_k$ by the submodules generated by $Q^{-D/2-k}$ is stationary, hence

$$Q^{-D/2-k-1} = P(t,\partial,D)\,Q^{-D/2-k}$$

for suitable choices of k and P. It remains to show that $\mathcal{P}$ is finitely generated and this is done using the filtration

$$\mathcal{P}_m = \{p(t,D)\,Q^{-D/2-m} \mid \deg p \leq m(\deg Q + 1)\},$$

which is compatible with the natural filtration of the ring of differential operators and whose growth

$$\dim \mathcal{P}_m \sim (\deg Q + 1)^n\, m^n$$

is slow enough to show that $\mathcal{P}$ is finitely generated (cf. [**17**]). □

Using Lemma 1.6 together with integration by parts and induction on the number of variables t_j, one can prove the following statement (cf. [**128**]).

LEMMA 1.7. *Let g be a smooth function of rapid decay defined for $t_j \geq 0$, whose derivatives are of rapid decay; then the integral*

$$I(D,g) = \int_{t_j \in [0,\infty)} g(t)\,Q(t)^{-D/2}\,dt$$

extends meromorphically to the whole complex plane.

PROOF. Indeed, integration by parts gives

$$I(D+2,g) = q(D)^{-1}\,I(D, L(D)^*\,g) + C(D), \tag{1.133}$$

where $L(D)^*$ is the adjoint of $L(D)$ and $C(D)$ is a boundary term which is meromorphic by the induction hypothesis. □

The problem is that, when one wants to apply this lemma to an expression of the form (1.132), the function which plays the role of g is given by

$$g(p,t) = e^{-\sum t_j m_j^2 - \sum t_j F_j(\kappa(t),p)^2}\,f(p,t). \tag{1.134}$$

This function, while of rapid decay along with its derivatives for t large, is not smooth for $t \to 0$. It is bounded and has a finite limit but its higher derivatives are in general not bounded and blow up by a negative power of $\det(A(t))$ so that there is no guarantee that integration by parts improves the domain of convergence. This seems to be a major difficulty, also witnessed in the notation of [**128**] by the presence of the term $-\mathrm{tr}(C\,B^{F^*}(t))$ in equation (34) of [**128**].

The problem disappears, however, if we look at the Taylor coefficients of the expansion of g at $p = 0$. Indeed, we have the following result.

LEMMA 1.8. *Let*

$$g(p,t) = \sum_\alpha g_\alpha(t)\, p^\alpha$$

be the Taylor expansion of $g(p,t)$ *at* $p = 0$. *Then all the functions* $g_\alpha(t)$ *are of the form* $Q(t)^{-|\alpha|} h_\alpha$ *where* h_α *fulfills the hypothesis of Lemma 1.7.*

PROOF. We just need to understand the behavior of the Taylor expansion at $p = 0$ of the expression

$$b(t,p) = e^{-\sum_j t_j F_j(\kappa(t,p),p)^2} = e^{a(t,p)},$$

where $\kappa(t,p)$ is the point where the function $\kappa \mapsto \sum_j t_j F_j(\kappa,p)^2$ reaches its minimum. For $p = 0$ this point is $\kappa(t,0) = 0$ so that $a(t,0) = 0$. Thus, each term in the Taylor expansion of $b(t,p)$ at $p = 0$ is a polynomial in the terms of the Taylor expansion of $a(t,p)$ at $p = 0$. Moreover, $\kappa(t,p)$ is of the form $A(t)^{-1}\ell(t,p)$ where $\ell(t,p)$ is linear in both t and p. This shows that, up to a power of $Q(t) = \det(A(t))$, all the terms in the Taylor expansion of both $a(t,p)$ and $b(t,p)$ at $p = 0$ are just polynomials in t. □

To summarize, we have obtained the following result.

THEOREM 1.9. *The Taylor coefficients at* $p = 0$ *of* $U^z(\Gamma(p_1,\ldots,p_N))$ *admit a meromorphic continuation to the whole complex plane* $z \in \mathbb{C}$.

The Laurent series expansion in the variable z around $z = 0$ for the Taylor coefficients at $p = 0$ of the expression $U^z(\Gamma(p_1,\ldots,p_N))$ serves as the basis for the *minimal subtraction* scheme which will be used in conjunction with DimReg in the next section.

DEFINITION 1.10. *Let* $L(z)$ *be the Laurent series in* z. *Let* $T(L(z))$ *denote its polar part. Minimal subtraction (MS) applied to* $L(z)$ *consists of the removal of the polar part* $L(z) - T(L(z))$.

In particular, one can apply this subtraction of the polar part to the Laurent series obtained by Theorem 1.9 for the Taylor coefficients at $p = 0$ of the expression $U^z(\Gamma(p_1,\ldots,p_N))$ of the unrenormalized values of Feynman graphs in dimensional regularization. In the following, with a commonly used abuse of terminology, we often refer loosely to the Laurent series expansion of $U^z(\Gamma(p_1,\ldots,p_N))$. It is understood that what we actually refer to is the result of Theorem 1.9 above.

It is important to remark at this point that, while subtraction of the pole part plays a key role in the *minimal subtraction* scheme, it is not sufficient to achieve renormalization. As we discuss in detail in Section 5 below, when the Feynman graphs become more complicated than our basic example (1.122), one encounters the problem of subdivergences and of the presence of *non-local* terms in the coefficients of the Laurent expansion. We shall see in the coming sections why this is a problem and how a renormalization procedure takes care of it.

Moreover, and this will play an important role later (cf. §6.6 below), expressions such as (1.132) have a physical dimension which depends on the parameter z. Indeed, at least at the formal level, the integrand $d^{D-z}k$ has the physical dimension of a mass to the power $D - z$. Since minimal subtraction involves subtracting the results for different values of z it is necessary to rescale them so that their physical dimensions no longer depend upon z. This is done by introducing an additional parameter μ, which has the physical units of a mass, and defining

$$(1.135)\qquad \begin{aligned} &U_\mu^z(\Gamma(p_1,\ldots,p_N)) = \\ &\int I_\Gamma(k_1,\ldots,k_L,p_1,\ldots,p_N)\mu^{zL}\, d^{D-z}k_1\cdots d^{D-z}k_L. \end{aligned}$$

This dependence on the mass (energy) scale plays an important role in renormalization.

For example, consider the case of the graph (1.122) in dimension $D = 6$. After dimensional regularization, one finds (cf. [**62**] p.173, where the symmetry factor $\sigma(\Gamma) = 2$ is also included) that the unrenormalized value is given by

$$(1.136)\qquad \begin{aligned} &U_\mu^z(\Gamma(p)) = \\ &\mu^z\,(4\pi)^{-3+z/2}\,g^2\,\Gamma(z/2-1)\,\int_0^1 (p^2(x-x^2)+m^2)^{1-z/2}\,dx. \end{aligned}$$

Before we begin discussing the issue of renormalization more closely, we should make a further remark about the *geometry* underlying the DimReg procedure. In fact, in setting (1.132) to be the "integral in dimension $D-z$", we have not actually made any attempt to define a space whose dimension is a complex number. In the physics literature, one takes (1.132) to be just a formal definition and regards the "complexified dimension" merely as a tool to perform calculations. In §19 of this Chapter we will see that there is more to DimReg than just a formal procedure. In fact, we will show a natural way to make sense of an actual space of dimension $D-z$, using noncommutative geometry.

5. The graph by graph method of Bogoliubov–Parasiuk–Hepp–Zimmermann

The basic idea of renormalization is that one eliminates the divergences step by step in the perturbative expansion by repeatedly adjusting the bare parameters in the Lagrangian so as to cancel the divergences. In the mass-shell scheme explained above this is achieved by setting the physical parameters, such as the mass or the coupling constant, to their observed values. This procedure quickly becomes very cumbersome and simpler renormalization schemes have been introduced in order to handle complicated theories such as the gauge theories involved in the Standard Model. Thus, in a chosen regularization scheme (for us DimReg+MS), one introduces a dependence on the regularization parameter (here the complex variable z) in the terms of the Lagrangian. One then adjusts the dependence of the bare parameters on z term by term in the perturbative expansion.

Let us consider, as a sufficiently generic example, a theory with Euclidean Lagrangian

$$(1.137)\qquad \mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 + \frac{m^2}{2}\phi^2 - \frac{g}{6}\phi^3,$$

where, as before, we changed the sign of the coupling constant to minimize the number of minus signs in the perturbative expansion.

We modify the parameters by introducing counterterms of the form

$$(1.138)\qquad \frac{1}{2}(\partial\phi)^2(1-\delta Z(z)) + \left(\frac{m^2-\delta m^2(z)}{2}\right)\phi^2 - \left(\frac{g+\delta g(z)}{6}\right)\phi^3.$$

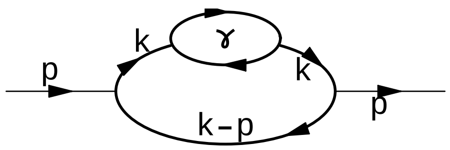


FIGURE 16. Subdivergence

These counterterms are divergent for $z \to 0$.

The main idea of the Bogoliubov–Parasiuk method is to write each counterterm itself as an infinite series

$$C(z) = \sum_{1\mathrm{PI}} \frac{C^z(\Gamma)}{\sigma(\Gamma)},$$

whose terms are labeled by suitable 1PI graphs Γ of the theory, while the value of the counterterm $C^z(\Gamma)$ is determined by a recursive procedure.
A theory is said to be *renormalizable* if it is possible to eliminate all divergences at all orders in the perturbative expansion by such a recursive procedure. The series of successive corrections to the bare parameters is typically itself divergent, but here one uses essentially the fact that the bare parameters in the Lagrangian are not physical observables that have to be finite, hence they can be modified in this way.

As we have seen, in order to be able to eliminate the divergences by adjusting the terms in the Lagrangian, it is necessary that the coefficients of the divergence in $U^z(\Gamma(p_1, \dots, p_N))$ be given by local terms, i.e. by a polynomial in the external momenta p_j, as only this type of term appears in the Lagrangian.

This is a necessary, not a sufficient condition, as one can see by considering the example of the integral (1.122) in different dimensions $D \in 4+2\mathbb{N}$. The coefficient of the pole is local, but for $D = 8$ it is of degree 4 and the theory is not renormalizable, since there is no term in the Lagrangian that would cancel it, while for $D = 6$ the pole coefficient has degree 2 and there are terms in the original Lagrangian $\mathcal{L}$ that can be used to eliminate the divergence by introducing suitable counterterms $\delta Z(z)$ and $\delta m^2(z)$.

As we mentioned in Section 4 above, the problem of non-local counterterms arises in conjunction with the presence of subdivergences in the Feynman graphs. We now show how this happens in a simple example.

5.1. The simplest example of subdivergence.

Let us consider the 1PI graph Γ (of the ϕ^3 theory) of Figure 16. To simplify the computation we work in the massless $D = 6$ case. Notice, however, that the general result of Theorem 1.9 above on the Taylor expansion at $p = 0$ works in the massive case. One has (up to a factor $g^4\mu^{2z}$)

$$U(\Gamma(p,-p)) = (2\pi)^{-2D} \int \frac{1}{k^4} \frac{1}{(k-p)^2} \frac{1}{(k+\ell)^2} \frac{1}{\ell^2} \, d^D\ell \, d^Dk. \tag{1.139}$$

We already computed the inner integral in ℓ and, disregarding the infrared divergence and setting $m = 0$ in (1.126), we get a multiplicative factor of

$$\int \frac{1}{(k+\ell)^2}\,\frac{1}{\ell^2}\,d^D\ell = \,k^{D-4}\,\pi^{D/2}\,\Gamma(2-D/2)\int_0^1 (x-x^2)^{D/2-2}\,dx. \tag{1.140}$$

One has

$$\int_0^1 (x-x^2)^{D/2-2}\,dx = \frac{\Gamma(D/2-1)^2}{\Gamma(D-2)},$$

whose poles at $D/2 - 1 \in -\mathbb{N}$ correspond to infrared divergences due to $m = 0$. Thus, after the integration in ℓ we are left with an integral which, up to the Γ factors and powers of π, is of the form

$$I = \int (k^2)^{D/2-4}\,\frac{1}{(k-p)^2}\,d^D k\,.$$

Writing then

$$x^{D/2-4} = \Gamma(4-D/2)^{-1}\int_0^\infty e^{-tx}\,t^{3-D/2}\,dt$$

one gets

$$I = \Gamma(4-D/2)^{-1}\int e^{-t_1k^2-t_2(k-p)^2}\,t_1^{3-D/2}\,dt_1\,dt_2\,d^D k.$$

With $t_1 = \lambda s$, $t_2 = \lambda(1-s)$, one has

$$t_1\,k^2 + t_2\,(k-p)^2 = \lambda\,q^2 + \lambda\,(s-s^2)p^2\,, \quad \text{with } \ q = k-(1-s)p.$$

Integration in $d^D q$ gives

$$I = \,\Gamma(4-D/2)^{-1}\,\pi^{D/2}\int e^{-\lambda\,(s-s^2)p^2}\,\lambda^{3-D}\,s^{3-D/2}\,\lambda d\lambda ds.$$

Integration in λ then gives the term $\Gamma(5-D)((s-s^2)p^2)^{D-5}$ and the remaining integration in s gives $\Gamma(D-4)\Gamma(D/2-1)\Gamma(3D/2-5)^{-1}$. Thus, one gets (up to a factor $g^4\mu^{2z}$)

$$U(\Gamma(p,-p)) = \,(4\pi)^{-D}\,\frac{\Gamma(2-\frac{D}{2})\Gamma(\frac{D}{2}-1)^3\Gamma(5-D)\Gamma(D-4)}{\Gamma(D-2)\Gamma(4-\frac{D}{2})\Gamma(\frac{3D}{2}-5)}\,(p^2)^{D-5}. \tag{1.141}$$

This gives a *double pole* at $D = 6$ coming from the factors $\Gamma(2-\frac{D}{2})\Gamma(5-D)$. If we expand in z at dimension $6-z$ and reinstate the μ dependence, the result comes from the expansion

$$(p^2/\mu^2)^{-z} = \sum \frac{(-z)^n}{n!}\,\log^n(p^2/\mu^2).$$

This contributes to the coefficient of $\frac{1}{z}$ by the non-polynomial term

$$-g^4\,(4\pi)^{-6}\,\frac{1}{18}\,p^2\,(\log(p^2/\mu^2)+\text{constant}). \tag{1.142}$$

The Fourier transform of the operator of multiplication by $p^2\,\log(p^2/\mu^2)$ is a non-local operator and no local counterterm in the Lagrangian can cancel this type of "non-local" divergence.

The conceptual origin of this type of non-polynomial term in the divergence can be understood by differentiating with respect to the external momentum p. In the case of the simple graph (1.122), differentiation with respect to p raises the term $(p+k)^2 + m^2$ in the denominator to higher powers, hence it quickly improves the convergence of the integral (1.122). This shows that the divergence is a polynomial in p. In the case discussed above, the part of the integral (1.139) which involves

$$\int \frac{1}{(k+\ell)^2}\,\frac{1}{\ell^2}\,d^D\ell \tag{1.143}$$

is independent of p and, irrespective of the repeated applications of ∂_p, it creates a term in $\frac{1}{z}$ in the overall integral. Using (1.140), one sees that in dimension $6-z$ the coefficient of $\frac{1}{z}$ in this term is given by

$$\int \frac{1}{(k+\ell)^2}\,\frac{1}{\ell^2}\,d^D\ell \sim -\frac{1}{3}\pi^3\,k^2\,\frac{1}{z} \tag{1.144}$$

and is local.

In other words, the effect of the *subdivergence* corresponding to the subgraph $\gamma \subset \Gamma$ is to add a divergent contribution to the integral (1.139). Regardless of the numerical coefficients involved in (1.139), this amounts to replacing the contribution of the expression (1.143) by the right-hand side of (1.144). Its effect on (1.139) is then the term

$$(2\pi)^{-2D}\int \frac{1}{k^4}\,\frac{1}{(k-p)^2}\,(-\frac{1}{3}\pi^3\,k^2\,\frac{1}{z})\,d^Dk, \tag{1.145}$$

which promotes the *finite* part of the graph (1.122)

$$\int \frac{1}{k^2}\,\frac{1}{(k-p)^2}\,d^Dk = -\frac{1}{3}\pi^3\,p^2\,\frac{1}{z} + \frac{1}{6}\pi^3\,p^2\,\log p^2 + \text{constant}$$

to a divergence and yields the non-local term (1.142).

We need to understand what happens in the method of differentiation with respect to the external momentum p after the replacement

$$\int \frac{1}{(k+\ell)^2}\,\frac{1}{\ell^2}\,d^D\ell \mapsto \int \frac{1}{(k+\ell)^2}\,\frac{1}{\ell^2}\,d^D\ell + \frac{1}{3}\pi^3\,k^2\,\frac{1}{z}. \tag{1.146}$$

The right-hand side is no longer divergent and is of the order of $k^2\log k^2$. This shows that, if we make the change (1.146) in the integral (1.139) and then differentiate enough times with respect to the external momentum p, we get a convergent integral. In other words, the correction (1.146) is good enough to restore the locality of the terms in the pole part.

The replacement (1.146) is expressed by the formula

$$\overline{R}(\Gamma) = U(\Gamma) + C(\gamma)U(\Gamma/\gamma) \tag{1.147}$$

where $\gamma \subset \Gamma$ is the divergent subgraph (Figure 16) and where the counterterm $C(\gamma)$ is given by the *opposite* of the pole part

$$C(\gamma) = -T(U(\gamma)). \tag{1.148}$$

The remaining graph Γ/γ is obtained by contracting γ to a single vertex. In (1.147) and in the following, we write $U(\Gamma)$ as a shorthand for $U(\Gamma(p_1,\ldots,p_N))$ where the external momenta p_i determine the external structure.

This *preparation* of the graph Γ by addition of counterterms corresponding to the subdivergence suffices to restore the locality of the divergence. In this simple massless example only one type of counterterm, namely the field strength counterterm, arises from subdivergences with two external lines. In the presence of a non-zero mass term there will also be a mass counterterm and we need to explain how (1.147) gets modified.

The contribution (1.140) becomes

$$\begin{aligned}\int \frac{1}{(k+\ell)^2+m^2}\,\frac{1}{\ell^2+m^2}\,d^D\ell = \\ \pi^{D/2}\,\Gamma(2-D/2)\int_0^1 ((x-x^2)k^2+m^2)^{D/2-2}\,dx.\end{aligned} \tag{1.149}$$

Notice that the coefficients of k^2 and m^2 in the divergence

$$(c_0 m^2 + c_1 k^2)\,\frac{1}{z} \tag{1.150}$$

in dimension $D = 6-z$ are different, $c_0 \neq c_1$, so that the corresponding counterterms need to be treated separately.

Thus, in order to prepare the graph Γ of Figure 16 in the massive case we need to modify (1.147) to the expression

$$\overline{R}(\Gamma) = U(\Gamma) + C(\gamma_{(0)})U(\Gamma/\gamma_{(0)}) + C(\gamma_{(1)})U(\Gamma/\gamma_{(1)}), \tag{1.151}$$

where now the additional index $j = 0, 1$ in $\gamma_{(j)}$ makes it possible to keep track of the distinction between the coefficients of k^2 and m^2 in the following way.

First, as explained in the discussion of the Feynman rules, we associate vertices of graphs to *all the terms* in the Lagrangian, including the kinetic term $\frac{1}{2}(\partial\phi)^2$ and the mass term $\frac{1}{2}m^2\phi^2$. Both of these vertices are of valence 2 (i.e. they correspond to quadratic terms), hence one needs a notation to distinguish between them, which we do by adding an index j at the vertex, with $j = 0$ for the mass term and $j = 1$ for the kinetic term. Thus, we get three types of vertices:

- Three-leg vertex $-\!\!<$ associated to the ϕ^3 term in the Lagrangian
- Two-leg vertex $\overset{0}{-\!\!\times\!\!-}$ associated to the term ϕ^2.
- Two-leg vertex $\overset{1}{-\!\!\times\!\!-}$ associated to the term $(\partial\phi)^2$.

When we evaluate a graph with such vertices we apply the Feynman rules with the coefficients corresponding to the terms of the initial Lagrangian as follows. The kinetic term gives a k^2, where k is the momentum flowing through the vertex, and the mass term gives an m^2. This is coherent with the overall homogeneity and the fact that, in the massless case, no mass counterterm is needed (in $D = 6$).

The role of the index j in the subgraphs $\gamma_{(j)}$ comes from their *external structure* determined by the external momenta p_i in $U(\Gamma(p_1, \ldots, p_N))$. The pole part (1.148) of this evaluation will (after preparation of γ) give a polynomial of the form $c_0 m^2 + c_1 k^2$ in terms of the external momentum k. We specify the external structure of $\gamma_{(j)}$ so that the evaluation of the graph with that structure gives the coefficient c_j. The coefficient c_0 is simply obtained by taking the external momenta to vanish and multiplying the result by m^{-2} so that the external structure corresponds to the distribution $m^{-2}\delta_0$, that is, to

$$\sigma_0(f) := m^{-2}\delta_0(f). \tag{1.152}$$

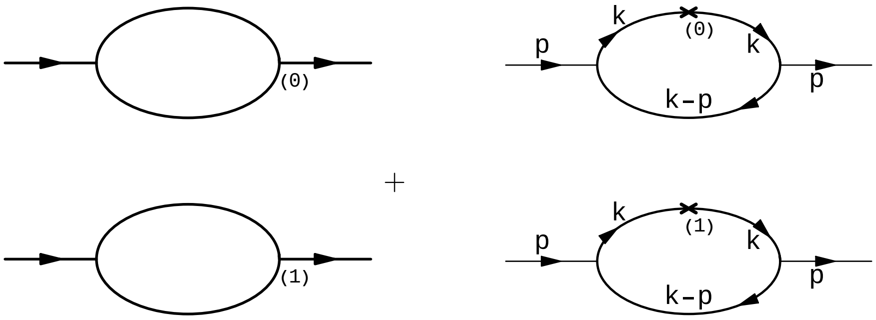


FIGURE 17. Preparation

The coefficient c_1 is obtained by differentiating with respect to k^2 i.e. by considering the distribution

$$\sigma_1(f) := (\frac{d^2}{dk^2}\delta_0)(f) = \frac{\partial^2}{\partial k^2} f(k)\bigg|_{k=0}. \tag{1.153}$$

Thus, in (1.151) the notation $\gamma_{(j)}$ denotes the graph γ with external structure σ_j.

Thus, the behavior of the counterterms $C(\gamma)$ is the reason why, when dealing with the external structure, one needs to assign two different distributions (1.152) and (1.153) that distinguish the vertices of valence two that come from the mass and from the kinetic terms in the free part of the Lagrangian.

In fact, as we just saw, the problem is that such a counterterm is not a scalar function but a combination $c_0 m^2 + c_1 k^2$. Thus, one needs to separate the counterterms $C(\gamma_{(i)})$ using any two distributions σ_i, $i = 0, 1$ with the property that

$$\sigma_0\,(c_0 m^2 + c_1 k^2) = c_0, \qquad \sigma_1\,(c_0 m^2 + c_1 k^2) = c_1\,. \tag{1.154}$$

Our choice (1.152), (1.153) clearly satisfies this requirement.

In the case of a massless theory, one does not take $k^2 = 0$ as in (1.152) and (1.153), to avoid a possible pole at $k = 0$ due to infrared divergences. The same procedure applies to give the counterterms, even though the finite part has a singularity.

Finally, notice that it is natural to take linear combinations of the values $U(\Gamma(p_1, \ldots, p_N))$ of a graph. The first example is the formula (1.95) which governs the contribution of Γ to the effective action.

5.2. Superficial degree of divergence.

One can get a rough idea of the behavior of the integral associated to a graph in D dimensions just by *power counting*, i.e. by assigning to each internal line a degree -2 corresponding to the homogeneity degree of the propagator, and to each integration variable a degree D. Thus, if we let

$$I = \#\Gamma^{(1)}_{\rm int}\,, \quad V = \#\Gamma^{(0)} \quad \text{and} \quad I - V = L - 1, \tag{1.155}$$

we get, by definition, the superficial degree of divergence

$$\deg(\Gamma) := -2I + DL. \tag{1.156}$$

When this degree is strictly negative it indicates that the only reason for a divergence of the graph comes from the presence of subdivergences.

Let us consider the ϕ^3 theory and compute this degree of graphs. We ignore the two-point vertices since they can only improve the convergence of the integral and also do not change the physical dimension of the result. Consider the set of pairs

$$\{(s,\ell) \in \Gamma^{(0)} \times \Gamma^{(1)} \mid s \in \ell\}. \tag{1.157}$$

Counting its cardinality using the two projections, one gets

$$3V = 2I + N, \tag{1.158}$$

where N is the number of external legs. It then follows that the degree of divergence can be expressed as

$$\deg(\Gamma) = 6 - 2N + (D-6)L. \tag{1.159}$$

This shows that the origin of the subdivergences is in subgraphs whose number of external legs is $N = 2$ or $N = 3$ (recall that tadpole graphs with $N = 1$ are excluded, cf. (1.56)).

5.3. Subdivergences and preparation.

The renormalization procedure is designed to solve two types of problems:

- *Organize the subdivergences*: when applying the regularization procedure to an unrenormalized value $U(\Gamma)$ one should be able to assume that all subgraphs $\gamma \subset \Gamma$ such that $U(\gamma)$ is itself divergent have already been regularized.
- *Eliminate the non-local terms*: The problem of non-local terms in the coefficients of the divergences of $U(\Gamma)$ arises in conjunction with the problem of subdivergences. The renormalization procedure should ensure that, when the latter is properly handled, one obtains only local terms as coefficients of the divergences.

A renormalization procedure that takes care of both problems was developed by Bogoliubov and Parasiuk [**26**] [**27**] and later refined by Hepp and Zimmermann [**166**], [**305**]. It is generally referred to as the BPHZ procedure.

We fix a renormalizable theory $\mathcal{T}$, and let J be the set of all monomials of the Lagrangian. It contains as a subset the set I of monomials of the interaction Lagrangian used in the definition of Feynman graphs as the target of the map (1.67). The definition of Feynman graphs extends with no change to graphs with vertices in J rather than I.

Definition 1.11. *Extended Feynman graphs are defined as in §3.3 above, with the map* (1.67) *replaced by*

$$\iota : \Gamma^{(0)} \to J, \tag{1.160}$$

subject to the same valence condition of (1.68).

We let $\mathrm{Graph}(\mathcal{T})$ be the set of extended Feynman graphs for the theory $\mathcal{T}$. The condition to be a graph in the sense of §3.3 is simply the requirement that the range of the map ι of (1.160) is contained in $I \subset J$.

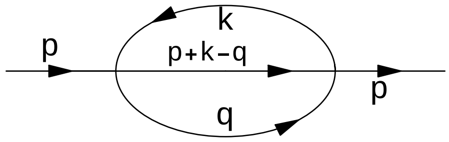


FIGURE 18. Graph Γ_{self} in ϕ^4, $\gamma = \{k, q\}$

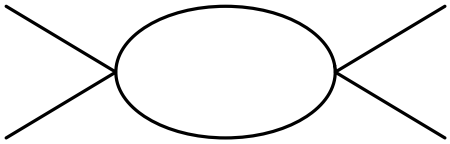


FIGURE 19. Subgraph of Γ_{self}

Let Γ be a graph of this theory. We need to specify what we mean by a subgraph of Γ and to concentrate on subgraphs that are responsible for subdivergences.

Let $\Gamma \in \text{Graph}(\mathcal{T})$, and consider a pair (γ, χ) where $\gamma \subset \Gamma^{(1)}_{\text{int}}$ is a subset of the set of internal lines of Γ and χ is a map from the set of connected components of $|\gamma| \subset |\Gamma|$ to the set J of monomials in the Lagrangian. For each connected component of $|\gamma|$ let $\gamma_i \subset \Gamma^{(1)}_{\text{int}}$ be the corresponding subset of edges and $\delta = \tilde{\gamma}_i$ be the graph, called a *component* of γ, defined as follows.

- The set of internal lines is $\delta^{(1)}_{\text{int}} = \gamma_i$.
- The set of vertices is $\delta^{(0)} = \Gamma^{(0)} \cap |\gamma_i|$.
- The set of external lines at a vertex v is the disjoint union of the $\partial_j^{-1}(v) \cap \gamma_i^c$, with the maps ∂_j as in (1.66).
- The maps ∂_j and ι of (1.66) and (1.67) are defined by restriction of the corresponding maps for Γ.

It is important to define the set of external lines as a disjoint union. This is illustrated in the example of Figure 24 (in the case of the ϕ^4 theory), where we take as the subgraph γ the subset of internal lines marked by the labels $\{2, 3\}$. It may happen that the graph $\tilde{\gamma}$ has more lines than Γ. This is shown in the example of the graph Γ of Figure 18 (again for the ϕ^4 theory), where we take for γ any subset $\gamma \subset \Gamma^{(1)}_{\text{int}}$ consisting of two elements. One obtains the graph $\tilde{\gamma}$ of Figure 19, which has 4 external lines and a total of six edges, while the original graph has only five.

Notice that, even though the γ_i are disjoint by construction, the components $\tilde{\gamma}_i$ need not be. In fact, they may have external edges in common. One can see this in the example of Figure 21.

Let (γ, χ) be as above. We let Γ/γ be the *contracted graph* obtained from the graph Γ by collapsing each component γ_i of $|\gamma|$ to a single vertex v_i. In other words, the contracted graph is obtained in the following way.

DEFINITION 1.12. *The contracted graph* Γ/γ *has the following description.*

- *The set of internal lines is* $\Gamma^{(1)}_{\text{int}} \setminus \gamma$.

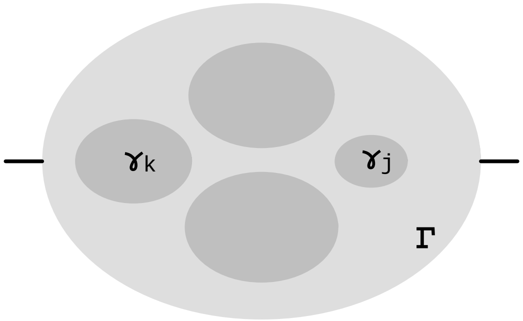


FIGURE 20. Components of $|\gamma|$

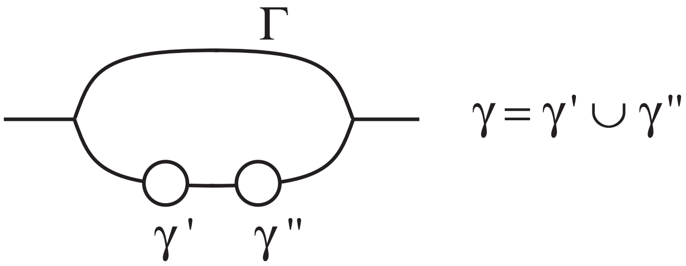


FIGURE 21. Components of $\gamma \in \mathcal{V}(\Gamma)$

- *The set of vertices is the disjoint union of the set* $\Gamma^{(0)} \setminus |\gamma|$ *and the set of connected components of* $|\gamma|$.
- *The set of external lines is the same as that of* Γ.
- *The maps* ∂_j *are defined by restriction of the corresponding maps for* Γ.
- *The map* ι *is given by restriction on* $\Gamma^{(0)}\setminus|\gamma|$ *and by* χ *on the set of connected components of* $|\gamma|$.

For instance, in the example of Figure 19 above, the contracted graph Γ/γ is given by Figure 22.

REMARK 1.13. It is not true in general that the construction of Definition 1.12 gives a graph which is also a Feynman graph of the theory, since the condition (1.68) is a strong restriction when applied to the vertices of Γ/γ coming from the components γ_i. Indeed, what this condition means is that, for each such component, the number of external lines of $\tilde{\gamma}_i$ is the degree of $\chi(\gamma_i)$.

For instance, in the case of the ϕ^3 theory, this means that the number of external lines of $\tilde{\gamma}_i$ is ≤ 3. For example, the subset γ of curved internal edges in Figure 23 does not qualify, since $\tilde{\gamma}$ has 4 external lines, so that the contracted graph would in fact be the ϕ^4 graph of Figure 22 which does not exist in ϕ^3 theory.

Thus, taking all this into account, we can now introduce the appropriate notion of subgraphs of a theory $\mathcal{T}$ in the following way.

DEFINITION 1.14. *Let* $\Gamma \in \mathrm{Graph}(\mathcal{T})$*; a* subgraph *of* Γ *is a pair* (γ, χ) *such that*

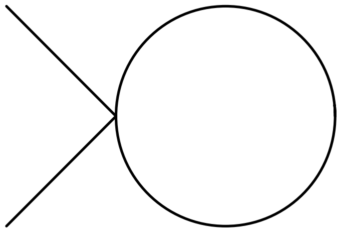


FIGURE 22. The contracted graph $\Gamma_{\mathrm{self}}/\gamma$ in ϕ^4

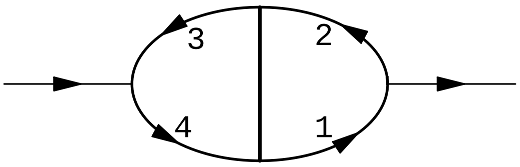


FIGURE 23. For $\gamma = \{1, 2, 3, 4\}$, $\tilde{\gamma}$ has 4 external lines

(1) *The components of* γ *are 1PI graphs.*
(2) $\Gamma/\gamma \in \mathrm{Graph}(\mathcal{T})$

We let $\mathcal{V}(\Gamma)$ *denote the set of (not necessarily connected) subgraphs of* Γ*, with* $\gamma \neq \Gamma$.

An example of the list of elements of $\mathcal{V}(\Gamma)$ with the corresponding $\tilde{\gamma}$ and Γ/γ is given for a ϕ^3 graph with three loops in Figure 25.

The subgraphs of Definition 1.14 and the corresponding contracted graphs satisfy the following general properties.

PROPOSITION 1.15. *1) Condition* (2) *of Definition* 1.14 *holds if and only if, for every component* γ_i *of* γ*, the degree of* $\chi(\gamma_i)$ *is equal to the number of external lines of* $\tilde{\gamma}_i$.
2) If Γ *is connected so is* Γ/γ.
3) Assume (1) *of Definition* 1.14. *Then* Γ *is a 1PI graph if and only if* Γ/γ *is 1PI.*

PROOF. 1) One needs to check that Γ/γ fulfills the condition of (1.68) at each vertex. For vertices v that are not in $|\gamma|$, the adjacent lines are not in γ, hence the condition follows since $\iota(v)$ is the same as for Γ. For a vertex v corresponding to a component γ_i of γ, the valence in Γ/γ is the number of external lines of $\tilde{\gamma}_i$, while by definition we have $\iota(v) = \chi(\gamma_i)$, so that the equivalence of the first statement follows.
Notice that, even if the component γ_i has only one vertex, the vertex of the graph $\tilde{\gamma}_i$ is not necessarily of the same valence. This can be seen by considering the two subgraphs (remember that one needs to specify χ so that there are two choices) of Figure 24 associated to the single line labeled 1. The original interaction vertex becomes a self-energy vertex.

2) There is a natural continuous surjection

$$\pi : |\Gamma| \to |\Gamma/\gamma| \tag{1.161}$$

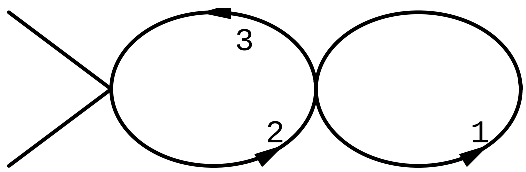


FIGURE 24. Two subgraphs correspond to $\{1\}$ and one to $\{2,3\}$ in ϕ^4

from the geometric realization of Γ to that of Γ/γ. Indeed, the latter is obtained by collapsing closed subsets to points. Thus, both 2) and the similar implication in 3) follow, since the image of a connected set by a continuous map is connected.

3) Let us prove the converse in 3). We need to show that, if we insert at a vertex v of a 1PI graph Γ_1 another 1PI graph Γ_2 with the right valence, we obtain a 1PI graph. Indeed, if one removes a line of the subgraph Γ_2, the latter remains connected and all its vertices are connected together. If one removes a line of Γ_1 the result is the remaining connected graph with Γ_2 inserted at a vertex, and this again is a connected graph. □

BPHZ procedure: We now describe the BPHZ procedure. One begins with a given Feynman graph Γ of the theory and performs the inductive *Bogoliubov–Parasiuk preparation*. This replaces the unrenormalized value $U(\Gamma)$ with the formal expression

$$\overline{R}(\Gamma) = U(\Gamma) + \sum_{\gamma\in\mathcal{V}(\Gamma)} C(\gamma)U(\Gamma/\gamma), \tag{1.162}$$

with $\mathcal{V}(\Gamma)$ as in Definition 1.14.

The counterterms $C(\gamma)$ in (1.162) are defined inductively, by setting

$$C(\Gamma) = -T(\overline{R}(\Gamma)) = -T\left(U(\Gamma) + \sum_{\gamma\in\mathcal{V}(\Gamma)} C(\gamma)U(\Gamma/\gamma)\right), \tag{1.163}$$

where T denotes the projection onto the pole part of the Laurent series in the variable z of DimReg. When γ is a union of different connected components, $\gamma = \gamma_1 \cup \cdots \cup \gamma_r$, the counterterm is *multiplicative*,

$$C(\gamma) = C(\tilde{\gamma}_1)\cdots C(\tilde{\gamma}_r), \tag{1.164}$$

with $\tilde{\gamma}$ the component graphs defined above (cf. [**62**] §5.3.3, p.105).

One defines then the *renormalized value* to be

$$R(\Gamma) = \overline{R}(\Gamma) + C(\Gamma) = U(\Gamma) + C(\Gamma) + \sum_{\gamma\in\mathcal{V}(\Gamma)} C(\gamma)U(\Gamma/\gamma), \tag{1.165}$$

where, as before, we have suppressed the explicit dependence on external momenta for simplicity of notation.

By the results of the previous sections, the expressions (1.162), (1.163), (1.165) are Laurent series in the variable z of DimReg, with the renormalized value $R(\Gamma)$ regular at $z = 0$ by construction (subtraction of the pole part).

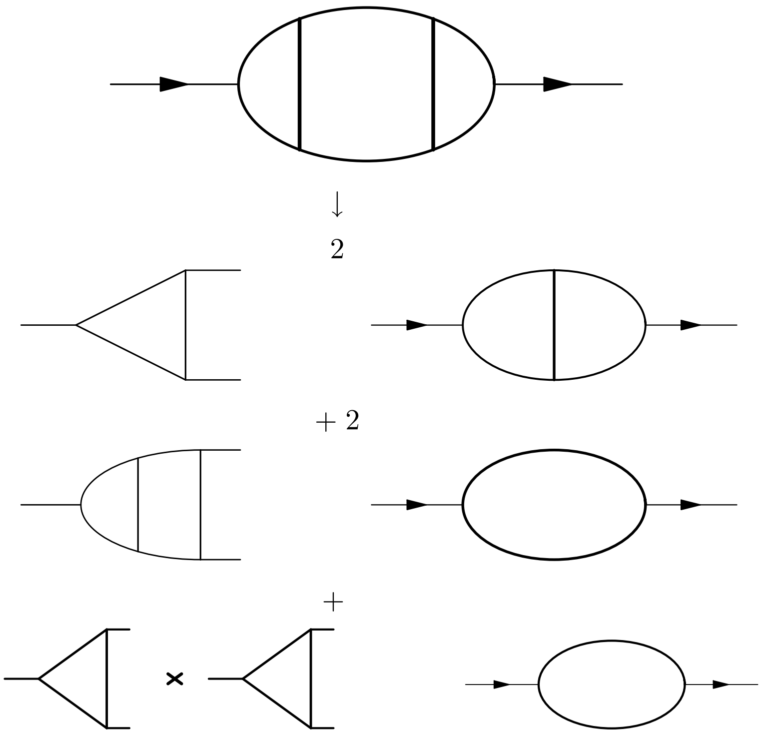


FIGURE 25. Subgraphs and the contracted graphs

REMARK 1.16. The external structure is defined for the component graphs by choosing distributions σ_ι for $\iota \in J$, which form a basis dual to the basis of monomials in the Lagrangian.

The only thing that matters is that we get the correct value of the counterterm, viewed as a function of the external momenta, when we sum all terms with the same number of external legs. Thus, in the ϕ^3 theory we have, for any two-leg graph,

$$C(\gamma)(p) = m^2\, C(\gamma_{(0)}) + p^2\, C(\gamma_{(1)}). \tag{1.166}$$

We also need to be more specific about the way the Feynman rules are extended to the graphs of Definition 1.11. Namely, we need to extend the Feynman rules to cover the vertices in $J \setminus I$, i.e. those vertices that do not belong to the original interaction Lagrangian and are used to handle the mass or field strength counterterms. The monomials in $J \setminus I$ correspond to the free part of the Euclidean action, hence they consist of terms of the form $\frac{1}{2}(\partial\phi)^2$ or $\frac{1}{2}m^2\phi^2$. We need to take into account an overall minus sign for the counterterms. In fact, these appear typically as $-C(\gamma)(p)$ in the Euclidean action. This means that, for each such vertex in $J \setminus I$, we apply the rules of §3.3 corresponding to the monomials $-\frac{1}{2}m^2\phi^2$ and $-\frac{1}{2}(\partial\phi)^2$. Thus, the vertices marked with $\times_{(0)}$ give the product of m^2 by a δ-function and the vertices marked with $\times_{(1)}$ give the product of k^2 (where k is the running momentum) by a δ-function. This has the advantage of minimizing the number of minus signs. Moreover, it also has the effect of dropping the new vertices entirely in the case of a massless theory. One should remember that only the product of the counterterms by these vertex contributions is relevant, so that one has the freedom of multiplying one by λ and the other by λ^{-1}. It is convenient to use from the start the negative sign for interaction terms (such as the term $-\frac{g}{6}\phi^3$) so that the minus sign of the counterterms is automatic for the monomials in I and the external structure of the component graphs is the evaluation at zero external momenta divided by the coupling constant.

The main result of [**26**], [**166**], [**305**] is the fact that this procedure takes care of both the problem of organizing subdivergences and of eliminating non-local terms. It is clear by construction that it accounts for subdivergences. Moreover, the following holds.

<u>Theorem</u> 1.17. (BPHZ) *1) The coefficients of the pole part of the prepared graph* $\overline{R}(\Gamma)$ *are given by local terms.*
2) The perturbative expansion of the functional integral with Euclidean Lagrangian

$$\mathcal{L}_E - \sum_{\iota(\Gamma^{(0)})\subset I} \frac{C(\Gamma)}{\sigma(\Gamma)} \tag{1.167}$$

gives the renormalized value of the theory.

To be more specific, in the case of the ϕ^3 theory, the above series takes the form

$$\begin{aligned} \tfrac{1}{2}\left(1 - \sum \frac{C(\Gamma_{(1)})}{\sigma(\Gamma)}\right)(\partial\phi)^2 \quad + \quad \tfrac{m^2}{2}\left(1 - \sum \frac{C(\Gamma_{(0)})}{\sigma(\Gamma)}\right)\phi^2 \\ - \quad \tfrac{g}{6}\left(1 + \sum \frac{C(\Gamma)}{\sigma(\Gamma)}\right)\phi^3, \end{aligned}$$

where in each case the graphs involved are those with $\iota(\Gamma^{(0)}) \subset I$.
Several of the following sections in this chapter are dedicated to explaining the rich mathematical structure that lies hidden behind the BPHZ formulae (1.162), (1.163), (1.165). Our presentation is based on the work of Connes–Kreimer [**82**], [**83**] and of the authors [**87**], [**89**].

6. The Connes–Kreimer theory of perturbative renormalization

The Connes–Kreimer (CK) theory provides a conceptual understanding of the BPHZ procedure in terms of the Birkhoff factorization of loops in a pro-unipotent complex Lie group associated to a commutative Hopf algebra of Feynman graphs. The main points, which we discuss in detail in the rest of this section, are summarized as follows.

- The Hopf algebra $\mathcal{H}$ of Feynman graphs.
- The BPHZ procedure as a Birkhoff factorization in the Lie group $G(\mathbb{C}) = \mathrm{Hom}(\mathcal{H}, \mathbb{C})$.
- The action of $G(\mathbb{C})$ on the coupling constants of the theory, through formal diffeomorphisms.
- The renormalization group as a 1-parameter subgroup of $G(\mathbb{C})$.

We begin by recalling some general facts about commutative Hopf algebras and affine group schemes. We then present the construction of the Hopf algebra $\mathcal{H}(\mathcal{T})$ of Feynman graphs of a renormalizable quantum field theory $\mathcal{T}$.

6.1. Commutative Hopf algebras and affine group schemes.

The theory of affine group schemes is developed in SGA 3 [**115**]. Whereas affine schemes are dual to commutative algebras, affine group schemes are dual to commutative Hopf algebras (for an introductory text see also [**290**]). We recall here some basic facts that we need later.

Definition 1.18. *Let k be a field of characteristic zero. A commutative Hopf algebra $\mathcal{H}$ over k is a commutative algebra with unit over k, endowed with a (not necessarily cocommutative) coproduct $\Delta : \mathcal{H} \to \mathcal{H} \otimes_k \mathcal{H}$, a counit $\varepsilon : \mathcal{H} \to k$, which are k-algebra morphisms and an antipode $S : \mathcal{H} \to \mathcal{H}$ which is a k-algebra antihomomorphism. These satisfy the "co-rules"*

$$
\begin{aligned}
&(\Delta \otimes id)\Delta = (id \otimes \Delta)\Delta && : \mathcal{H} \to \mathcal{H} \otimes_k \mathcal{H} \otimes_k \mathcal{H}, \\
&(id \otimes \varepsilon)\Delta = id = (\varepsilon \otimes id)\Delta && : \mathcal{H} \to \mathcal{H}, \\
&m(id \otimes S)\Delta = m(S \otimes id)\Delta = 1\,\varepsilon && : \mathcal{H} \to \mathcal{H},
\end{aligned}
\tag{1.168}
$$

where we use m to denote multiplication in $\mathcal{H}$.

Suppose given a commutative Hopf algebra $\mathcal{H}$ as above. There is an associated covariant functor G from the category $\mathcal{A}_k$ of unital k-algebras to the category $\mathcal{G}$ of groups. It assigns to a unital k-algebra A the group

$$G(A) = \mathrm{Hom}_{\mathcal{A}_k}(\mathcal{H}, A). \tag{1.169}$$

Thus, elements of the group $G(A)$ are k-algebra homomorphisms

$$\phi : \mathcal{H} \to A\,, \quad \phi(x\,y) = \phi(x)\phi(y), \quad \forall x, y \in \mathcal{H}\,, \quad \phi(1) = 1\,.$$

The product in $G(A)$ is dual to the coproduct of $\mathcal{H}$, that is,

$$\phi_1 * \phi_2\,(x) = \langle \phi_1 \otimes \phi_2\,,\ \Delta(x) \rangle\,. \tag{1.170}$$

Similarly, the inverse and the unit of $G(A)$ are determined by the antipode and the co-unit of $\mathcal{H}$. The co-rules imply that these operations define a group structure on $G(A)$. The functor G is an *affine group scheme.*

One can give the following general definition.

Definition 1.19. *An affine group scheme G is a covariant representable functor from the category of commutative algebras over k to the category of groups.*

In fact, the functor G of (1.169) is certainly representable (by $\mathcal{H}$) and, conversely, any covariant representable functor from the category of commutative algebras over k to groups is an affine group scheme G, represented by a commutative Hopf algebra, uniquely determined up to canonical isomorphism (cf. e.g. [**290**]).

Here are some simple examples of affine group schemes.

- The additive group $G = \mathbb{G}_a$ corresponds to the Hopf algebra $\mathcal{H} = k[t]$ with coproduct $\Delta(t) = t \otimes 1 + 1 \otimes t$.
- The multiplicative group $G = \mathbb{G}_m$ is the affine group scheme obtained from the Hopf algebra $\mathcal{H} = k[t, t^{-1}]$ with coproduct $\Delta(t) = t \otimes t$.

- The group of roots of unity μ_n is the kernel of the homomorphism $\mathbb{G}_m \to \mathbb{G}_m$ given by raising to the n-th power. It corresponds to the Hopf algebra $\mathcal{H} = k[t]/(t^n - 1)$.
- The affine group scheme $G = \mathrm{GL}_n$ corresponds to the Hopf algebra

$$\mathcal{H} = k[x_{i,j}, t]_{i,j=1,\dots,n}/(\det(x_{i,j})t - 1),$$

with coproduct $\Delta(x_{i,j}) = \sum_k x_{i,k} \otimes x_{k,j}$.

The latter example is quite general. In fact, if $\mathcal{H}$ is finitely generated as an algebra over k, then the corresponding affine group scheme G is a linear algebraic group over k, and can be embedded as a Zariski closed subset in some GL_n.

More generally, we have the following result (cf. e.g. [**231**] Proposition 4.13 and [**290**]).

LEMMA 1.20. *Let $\mathcal{H}$ be a commutative positively graded connected Hopf algebra. There exists a family $\mathcal{H}_i \subset \mathcal{H}$, $i \in \mathcal{I}$, indexed by a partially ordered set, where the $\mathcal{H}_i$ are finitely generated algebras over k satisfying the following properties:*

(1) *$\Delta(\mathcal{H}_i) \subset \mathcal{H}_i \otimes \mathcal{H}_i$, for all $i \in \mathcal{I}$.*
(2) *$S(\mathcal{H}_i) \subset \mathcal{H}_i$, for all $i \in \mathcal{I}$.*
(3) *For all $i, j \in \mathcal{I}$, there exists a $k \in \mathcal{I}$ such that $\mathcal{H}_i \cup \mathcal{H}_j \subset \mathcal{H}_k$, and $\mathcal{H} = \cup_i \mathcal{H}_i$.*

Then the affine group scheme G of $\mathcal{H}$ is of the form

$$G = \varprojlim_i G_i, \tag{1.171}$$

where the G_i are the linear algebraic groups dual to $\mathcal{H}_i$.

Thus, in general, such an affine group scheme is a projective limit of unipotent linear algebraic groups G_i, we say that G is a pro-unipotent affine group scheme.

Recall that an element X in a Hopf algebra $\mathcal{H}$ is said to be group-like if $\Delta(X) = X \otimes X$ and is said to be primitive if $\Delta(X) = X \otimes 1 + 1 \otimes X$. If G denotes the affine group scheme of a commutative Hopf algebra $\mathcal{H}$, then a group-like element $X \in \mathcal{H}$ determines a homomorphism $G \to \mathbb{G}_m$, which, at the level of Hopf algebras, is given by $\phi : k[t, t^{-1}] \to \mathcal{H}$ with $\phi(t) = X$. Similarly, a primitive element $X \in \mathcal{H}$ corresponds to a homomorphism $G \to \mathbb{G}_a$.

There is a notion of Lie algebra for an affine group scheme. It is also defined as a functor.

DEFINITION 1.21. *The Lie algebra of an affine group scheme G is a covariant functor $\mathfrak{g} = \mathrm{Lie}\, G$ from the category $\mathcal{A}_k$ of commutative k-algebras to the category $\mathcal{L}_k$ of Lie algebras over k,*

$$A \mapsto \mathfrak{g}(A)\,, \tag{1.172}$$

where $\mathfrak{g}(A)$ is the Lie algebra of linear maps $L : \mathcal{H} \to A$ satisfying

$$L(X\,Y) = L(X)\,\varepsilon(Y) + \varepsilon(X)\,L(Y)\,, \quad \forall X\,, Y \in \mathcal{H}\,, \tag{1.173}$$

where ε is the counit of $\mathcal{H}$.

The Lie bracket of two elements in $\mathfrak{g}(A)$ is given by

$$[L_1,\, L_2](X) = \langle L_1 \otimes L_2 - L_2 \otimes L_1, \Delta(X)\rangle\,. \tag{1.174}$$

Normally, the datum $\mathrm{Lie}\, G$ of the Lie algebra is not sufficient to reconstruct the affine group scheme G. One can see this already in the simplest case: the affine group schemes $\mathbb{G}_m$ and $\mathbb{G}_a$ have the same Lie algebra. There is, however, a particular class of cases in which the knowledge of the Lie algebra is sufficient to reconstruct the Hopf algebra. This follows from the Milnor–Moore Theorem [**231**] Theorem 5.18, which asserts that a primitively generated Hopf algebra over a field k of characteristic zero is the enveloping algebra of the Lie algebra of primitive elements.
Recall that, given a graded connected Hopf algebra $\mathcal{H}$, one obtains a dual Hopf algebra $\mathcal{H}^\vee$ by reversing all the arrows. More precisely we assume that the graded pieces $\mathcal{H}_n$ are finite-dimensional k-vector spaces, and we define $\mathcal{H}^\vee$ as the graded dual of $\mathcal{H}$, i.e. an element of $\mathcal{H}^\vee$ is a finite linear combination of homogeneous linear forms on $\mathcal{H}$. It follows from [**231**] Proposition 3.1, that $\mathcal{H}^\vee$ is a graded Hopf algebra, and also that the bidual $(\mathcal{H}^\vee)^\vee$ is canonically isomorphic to $\mathcal{H}$. We can now state the following corollary of the Milnor–Moore Theorem:

THEOREM 1.22. *Let $\mathcal{H}$ be a commutative Hopf algebra over a field k of characteristic zero. Assume that $\mathcal{H}$ is positively graded and connected, $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$, with $\mathcal{H}_0 = k$, and that the graded pieces $\mathcal{H}_n$ are finite-dimensional k-vector spaces. Let $\mathcal{H}^\vee$ be the dual Hopf algebra and let $\mathcal{L}$ denote the Lie algebra of primitive elements in $\mathcal{H}^\vee$. Then there is a canonical isomorphism of Hopf algebras*

$$\mathcal{H} \cong U(\mathcal{L})^\vee, \tag{1.175}$$

where $U(\mathcal{L})$ is the universal enveloping algebra of $\mathcal{L}$. Moreover, $\mathcal{L} = \mathrm{Lie}\, G(k)$ as a graded Lie algebra.

PROOF. Let $\mathcal{H}^\vee$ be the graded dual of $\mathcal{H}$. By [**231**] Corollary 4.18, $\mathcal{H}^\vee$ is primitively generated. Thus the Milnor–Moore Theorem ([**231**] Theorem 5.18) shows that $\mathcal{H}^\vee \cong U(\mathcal{L})$. The result follows from the isomorphism $\mathcal{H} \cong (\mathcal{H}^\vee)^\vee$. □

REMARK 1.23. Note that without the finiteness hypothesis, the dual $\mathcal{H}^\vee$ of a Hopf algebra $\mathcal{H}$ is ill defined since, if one considers arbitrary linear forms L on $\mathcal{H}$, there is no guarantee that the functional $X \otimes Y \mapsto L(XY)$ can be expressed as a finite linear combination of simple tensors $L_1 \otimes L_2$. This difficulty disappears if one restricts to linear forms for which this finiteness holds (at any level), and this is automatic for primitive elements, as they fulfill by hypothesis the equality

$$L(XY) = L(X)\,\varepsilon(Y) + \varepsilon(X)\,L(Y)\,, \quad \forall X, Y \in \mathcal{H}\,.$$

REMARK 1.24. In our set-up we are only considering Hopf algebras which are commutative rather than "graded-commutative" as in [**231**]. This should not create confusion since one can multiply the grading by 2 to apply the results of [**231**]. In fact, in our setting one should really think of the grading as being an even grading: this will become clearer in comparison with the Lie algebras of motivic Galois groups of mixed Tate motives (see Section 8), where the grading corresponds to the weight filtration, which is naturally parameterized by even integers. A reason for regarding the grading by number of internal lines of graphs (cf. Proposition 1.30 below) as an *even* grading comes from the fact that it is customary in physics to think of internal lines of graphs as a *pair* of half-lines, see for instance [**170**], [**19**]. In fact we saw in §3.1 that the graphs are obtained from pairings of half-lines. Thus the grading is actually given by the number of half-lines that contribute to the internal lines of the graph. In the mathematical literature, this formulation of graphs in terms of collections of half-lines was variously used (cf. [**189**], [**142**]).

The result of Theorem 1.22 can also be used to obtain an explicit description of the Lie algebra of an affine group scheme G, using the primitive elements in the dual Hopf algebra. This is the form in which it is used in the Connes-Kreimer theory.

LEMMA 1.25. *Let $\mathcal{H}$ be a commutative graded connected Hopf algebra, $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$. There is an action of the multiplicative group $\mathbb{G}_m$ on $\mathcal{H}$ and an associated semidirect product of affine group schemes $G^* = G \rtimes \mathbb{G}_m$.*

PROOF. Let Y denote the generator of the grading, namely the linear operator on $\mathcal{H}$ that satisfies $Y(X) = nX$ for all $X \in \mathcal{H}_n$. This defines an action of the multiplicative group $\mathbb{G}_m$ on $\mathcal{H}$ by setting, for all $u \in \mathbb{G}_m$,

$$u^Y(X) = u^n X \quad \forall X \in \mathcal{H}_n. \tag{1.176}$$

Thus, we can consider the affine group scheme obtained as the semidirect product

$$G^* = G \rtimes \mathbb{G}_m, \tag{1.177}$$

and we have a corresponding homomorphism $G^* \to \mathbb{G}_m$. The action (1.176) is also defined on the dual Hopf algebra $\mathcal{H}^\vee$. One then obtains an explicit description of the Lie algebra of G^* in terms of the Lie algebra of G. Namely, the Lie algebra of G^* has an additional generator Z_0 such that

$$[Z_0, X] = Y(X) \quad \forall\, X \in \mathrm{Lie}\, G\,. \tag{1.178}$$

□

6.2. The Hopf algebra of Feynman graphs: discrete part.

The first main step of the CK theory is to associate to a given renormalizable quantum field theory $\mathcal{T}$ a Hopf algebra $\mathcal{H}(\mathcal{T})$ over $k = \mathbb{C}$, where the coproduct reflects the combinatorics of the BPHZ preparation formula (1.162). In this section we describe a simplified version of the Hopf algebra of Feynman graphs (the "discrete part") where we only consider graphs responsible for divergences. We discuss in §6.3 the full Hopf algebra of Feynman graphs, where we take into account arbitrary 1PI graphs. We let as above J denote the set of all monomials in the Lagrangian.

DEFINITION 1.26. *For a given renormalizable quantum field theory $\mathcal{T}$, the* discrete Hopf algebra of Feynman graphs *$\mathcal{H}(\mathcal{T})$ is the free commutative algebra over $\mathbb{C}$ generated by pairs (Γ, w) with $\Gamma \in \mathrm{Graph}(\mathcal{T})$ a 1PI graph, and $w \in J$ a monomial with degree equal to the number of external lines of Γ. The coproduct is defined on generators as*

$$\Delta(\Gamma) = \Gamma \otimes 1 + 1 \otimes \Gamma + \sum_{\gamma \in \mathcal{V}(\Gamma)} \gamma \otimes \Gamma/\gamma, \tag{1.179}$$

where the set $\mathcal{V}(\Gamma)$ is as in Definition 1.14.

Notice that by construction the meaning of γ in (1.179) is the product of the components $\tilde{\gamma}_i$ where each is endowed with the element $\chi(\gamma_i) \in J$. In other words,

$$\gamma = \prod (\tilde{\gamma}_i, \chi(\gamma_i)).$$

The external lines of Γ/γ are the same as for Γ, hence in (1.179) the notation Γ/γ stands for the pair $(\Gamma/\gamma, w)$ with the same w as for Γ.
Notice how $\mathcal{H}(\mathcal{T})$ is strongly dependent on the physical theory $\mathcal{T}$, both in the generators and in the structure of the coproduct, where the class of subgraphs $\mathcal{V}(\Gamma)$ also depends on $\mathcal{T}$. We see in §7.6 below that there is a universal Hopf algebra that relates naturally to all the $\mathcal{H}(\mathcal{T})$ and encodes the renormalization procedure canonically for all the physical theories.

THEOREM 1.27. ([**82**]) *The commutative algebra of Definition* 1.26 *with the coproduct* (1.179) *is a Hopf algebra* $\mathcal{H}(\mathcal{T})$.

PROOF. Let us prove that Δ is coassociative, i.e. that

$$(\Delta \otimes \mathrm{id})\,\Delta = (\mathrm{id} \otimes \Delta)\,\Delta\,. \tag{1.180}$$

Since both sides of (1.180) are algebra homomorphisms from $\mathcal{H}$ to $\mathcal{H}\otimes\mathcal{H}\otimes\mathcal{H}$, it will be enough to check that they give the same result on 1PI graphs Γ. It is suitable to enlarge the definition of $\mathcal{V}(\Gamma)$ to $\bar{\mathcal{V}}(\Gamma)$ which includes the cases $\gamma = \emptyset$ and $\gamma = \Gamma^{(1)}_{\rm int}$, allowing the graphs $\tilde{\gamma}$ or Γ/γ to be empty, in which case we represent them by the symbol $1 \in \mathcal{H}(\mathcal{T})$. We write

$$\gamma \preceq \Gamma \Leftrightarrow \gamma \in \bar{\mathcal{V}}(\Gamma) \tag{1.181}$$

and

$$\gamma \prec \Gamma \Leftrightarrow \gamma \in \bar{\mathcal{V}}(\Gamma)\,, \gamma \neq \Gamma^{(1)}_{\rm int}. \tag{1.182}$$

With this notation, the coproduct of Γ takes the form

$$\Delta\,\Gamma = \sum_{\gamma\preceq\Gamma} \tilde{\gamma} \otimes \Gamma/\gamma, \tag{1.183}$$

where by definition $\tilde{\gamma}$ is given by the product of its components

$$\tilde{\gamma} = \prod \tilde{\gamma}_j. \tag{1.184}$$

Applying $\Delta \otimes \mathrm{id}$ on both sides of (1.183) we get

$$(\Delta \otimes \mathrm{id})\,\Delta\,\Gamma = \sum_{\gamma\preceq\Gamma} \Delta\,\tilde{\gamma} \otimes \Gamma/\gamma. \tag{1.185}$$

Since Δ is a homomorphism, one has

$$\Delta(\tilde{\gamma}) = \prod \Delta(\tilde{\gamma}_j).$$

Using (1.183), this can be written as

$$\Delta(\tilde{\gamma}) = \sum_{\gamma'\preceq\tilde{\gamma}} \tilde{\gamma}' \otimes \tilde{\gamma}/\gamma', \tag{1.186}$$

where the notation $\preceq$ has been extended to non-connected graphs. What matters here is that a subgraph of $\tilde{\gamma}$ is given by a collection of subgraphs of its components, with the limit cases in $\bar{\mathcal{V}}$ allowed. Since things happen independently in each component, the product of the sums is equal to the sum of the products and one gets (1.186). We can think of γ' as a collection of subsets γ'_j of the sets of internal lines of the components $\tilde{\gamma}_j$. Thus, we can view the union of the γ'_j as a subset γ' of the set of internal lines of Γ. Let us check that, for $\gamma' \subset \gamma$, we have

$$\gamma' \in \bar{\mathcal{V}}(\Gamma) \Leftrightarrow \gamma'_j \in \bar{\mathcal{V}}(\tilde{\gamma}_j)\,, \quad \forall j. \tag{1.187}$$

Here, for each component γ_j of γ, one lets $\gamma'_j = \gamma_j \cap \gamma'$. First, a connected component δ of γ' is contained in one (and only one) of the γ_j. Let us show that the graph $\tilde{\delta}$ is the same relative to $\tilde{\gamma}_j$ or to Γ. Its set of internal lines is δ in both cases. Its set of vertices is the set of vertices of $|\delta|$, which does not change in passing from $\tilde{\gamma}_j$ to Γ. The set of external lines at a vertex v is the disjoint union of the $\partial_i^{-1}(v) \cap \delta^c$. One has to show that this is unchanged in passing from $\tilde{\gamma}_j$ to Γ. Since a line ℓ of Γ such that $\partial_i \, \ell = v$ is a line of $\tilde{\gamma}_j$, one gets the result. It follows that the graph $\tilde{\delta}$ does not depend on whether it is taken relative to $\tilde{\gamma}_j$ or to Γ and we can now write

$$(\Delta \otimes \mathrm{id}) \, \Delta \, \Gamma = \sum_{\gamma' \preceq \gamma \preceq \Gamma} \tilde{\gamma}' \otimes \tilde{\gamma}/\gamma' \otimes \Gamma/\gamma, \tag{1.188}$$

where the parameters γ and γ' vary among the subsets of $\Gamma^{(1)}_{\mathrm{int}}$ that fulfill

$$\gamma \in \bar{\mathcal{V}}(\Gamma) \,, \quad \gamma' \in \bar{\mathcal{V}}(\Gamma) \,, \quad \gamma' \subset \gamma. \tag{1.189}$$

We can write (using (1.183)) the coproduct of Γ as

$$\Delta \, \Gamma = \sum_{\gamma' \preceq \Gamma} \tilde{\gamma}' \otimes \Gamma/\gamma', \tag{1.190}$$

hence

$$(\mathrm{id} \otimes \Delta) \, \Delta \, \Gamma = \sum_{\gamma' \preceq \Gamma} \tilde{\gamma}' \otimes \Delta \, (\Gamma/\gamma'). \tag{1.191}$$

Thus, to prove (1.180) it suffices to show that, for fixed $\gamma' \in \bar{\mathcal{V}}(\Gamma)$, one has

$$\Delta \, (\Gamma/\gamma') = \sum_{\gamma \preceq \Gamma, \, \gamma \supset \gamma'} \tilde{\gamma}/\gamma' \otimes \Gamma/\gamma. \tag{1.192}$$

We let $\Gamma' = \Gamma/\gamma'$ be the contracted graph. The set of internal lines of Γ' is the complement of γ' in $\Gamma^{(1)}_{\mathrm{int}}$. Thus, subsets $\gamma \supset \gamma'$ correspond to subsets of $(\Gamma')^{(1)}_{\mathrm{int}}$ by the map $\gamma \mapsto \rho(\gamma) = \gamma'' = \gamma \backslash \gamma'$. One has (using (1.183)),

$$\Delta \, \Gamma' = \sum_{\gamma'' \preceq \Gamma'} \tilde{\gamma}'' \otimes \Gamma'/\gamma''. \tag{1.193}$$

It remains to show that the map ρ fulfills

$$\gamma \preceq \Gamma \Leftrightarrow \rho(\gamma) \preceq \Gamma' \,, \quad \tilde{\gamma}/\gamma' = \widetilde{\rho(\gamma)} \,, \quad \Gamma/\gamma = \Gamma'/\rho(\gamma). \tag{1.194}$$

We take a subset $\gamma \supset \gamma'$ and we let

$$\pi : |\Gamma| \to |\Gamma/\gamma'| \tag{1.195}$$

be the continuous surjection of (1.161). It is not injective but, if $\pi(x) = \pi(y)$, then x and y are in the same connected component of $|\gamma'|$. Since we have $\gamma' \subset \gamma$, each connected component of $|\gamma'|$ is inside a component of $|\gamma|$. Thus, if $\pi(x) = \pi(y)$, then they are in the same connected component of $|\gamma|$. We let γ_i be the components of γ. The argument above shows that the $\pi(|\gamma_i|)$ are disjoint. The lines of $\pi(|\gamma_i|)$ are labelled by the complement $\gamma''_i = \gamma_i \backslash \gamma'$ of $\gamma' \cap \gamma_i$ in γ_i. Thus, γ''_i is a connected component of $|\gamma''|$, where $\rho(\gamma) = \gamma'' = \gamma \backslash \gamma'$.

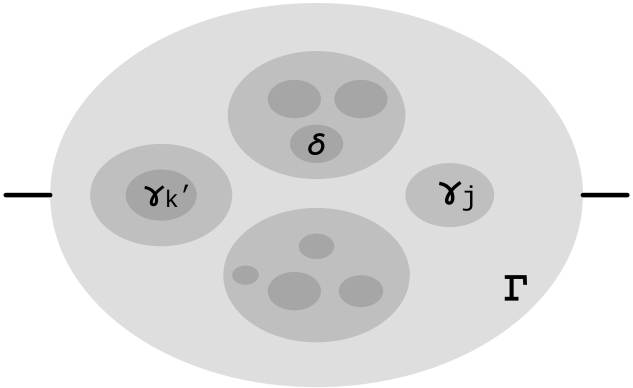


FIGURE 26. The pair $\gamma' \subset \gamma$, components δ of γ' are the darker ones

Conversely, every connected component $|\gamma''_j|$ of $|\gamma''|$ is obtained in this way and its inverse image $\pi^{-1}(|\gamma''_j|)$ is connected and is a component $|\gamma_i|$ of $|\gamma|$. The corresponding graph $\tilde{\gamma}_i$ is obtained by inserting at some vertices of $\tilde{\gamma}''_j$ the corresponding 1PI graphs associated to the components of γ'. This operation preserves the property of being 1PI by Proposition 1.15. Moreover, the collapsed graph of a 1PI graph is still 1PI. This makes it possible to check, using Definition 1.14, that $\gamma \preceq \Gamma \Leftrightarrow \rho(\gamma) \preceq \Gamma'$. Moreover, the connected components of $|\gamma|$ and $|\gamma''|$ correspond bijectively under π. For each component γ_i of γ, one has $\tilde{\gamma}_i/\gamma' = \tilde{\gamma}''_i$ with $\gamma''_i = \gamma_i \backslash \gamma'$. To check this one needs to show that the external lines correspond. This follows from the equality $\gamma = \gamma' \cup \gamma''$.

One then needs to show that $\Gamma'/\gamma'' = \Gamma/\gamma$, that is, that $(\Gamma/\gamma')/\gamma'' = \Gamma/\gamma$. In both cases the set of internal lines is the complement of $\gamma = \gamma' \cup \gamma''$ in $\Gamma^{(1)}_{\rm int}$. Let us look at vertices. The set of vertices of Γ/γ is the disjoint union of $\Gamma^{(0)} \backslash \gamma$ with the set of connected components γ_i of γ. In other words, it is the quotient of $\Gamma^{(0)}$ by the identification of vertices in the same component γ_i of γ. In the same way the set of vertices of $(\Gamma/\gamma')/\gamma''$ is obtained as a quotient of $\Gamma^{(0)}$. One first identifies vertices in the same component of γ' and then of γ''. Under the above correspondence between components γ_i of γ and components of γ'', one gets the desired identification of the sets of vertices. The external legs are unaltered in the collapsing process, hence one gets $\Gamma/\gamma = \Gamma'/\rho(\gamma)$ as required.

We have neglected to take the map χ into account in the above discussion and we explain how to handle it now. Indeed, a subgraph is not just given as a subset $\gamma \subset \Gamma^{(1)}_{\rm int}$ but it also consists of the datum of a map χ from the set of connected components of $|\gamma|$ to J. All the sums considered here above involve summations over compatible choices of χ. Thus, in (1.188) one has to sum over all choices of a χ for γ and a χ'_i for the $\gamma'_i \subset \gamma_i$, where the γ_i are the components of γ. The only requirement is that, if $\gamma'_i = \gamma_i$, then $\chi'_i(\gamma'_i) = \chi(\gamma_i)$. In any case, giving χ and the family χ'_i is the same as giving a χ' for the subset $\gamma' \subset \Gamma^{(1)}_{\rm int}$, and a χ'' for the subgraph γ/γ' of Γ'. Thus, the sums involved in (1.188) and (1.191) give the same result.

This ends the proof of the coassociativity. One still needs to show the existence of the antipode for $\mathcal{H}(\mathcal{T})$. This can be obtained ([**82**]) by an inductive procedure. In fact, this follows from the existence (cf. Proposition 1.30) of gradings of $\mathcal{H}(\mathcal{T})$ given by maps from Graph$(\mathcal{T})$ to $\mathbb{N}^*$ such that

$$\sum \delta(\gamma_i) + \delta(\Gamma/\gamma) = \delta(\Gamma)\,, \quad \forall \gamma \in \mathcal{V}(\Gamma). \tag{1.196}$$

This implies that, for a monomial $X \in \mathcal{H}(\mathcal{T})$, the coproduct $\Delta(X)$ can be written in the form

$$\Delta(X) = X \otimes 1 + 1 \otimes X + \sum X' \otimes X'', \tag{1.197}$$

where the terms X' and X'' are of strictly lower degree in the chosen grading satisfying (1.196). The equation $m(S \otimes 1)\Delta = \varepsilon 1$ can then be solved inductively for the antipode S using the formula

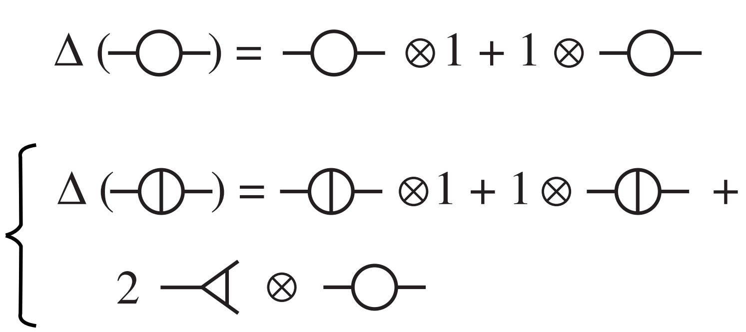
$$S(X) = -X - \sum S(X')X''. \tag{1.198}$$

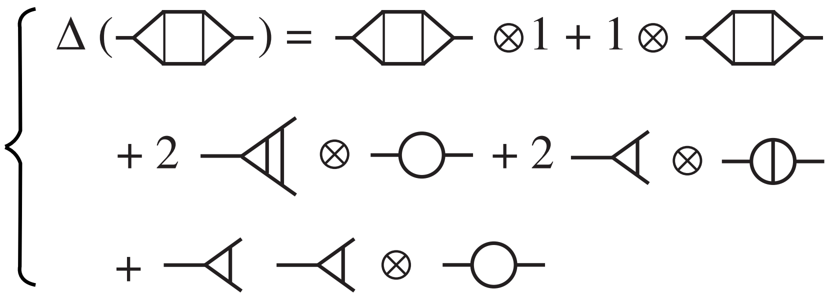
□

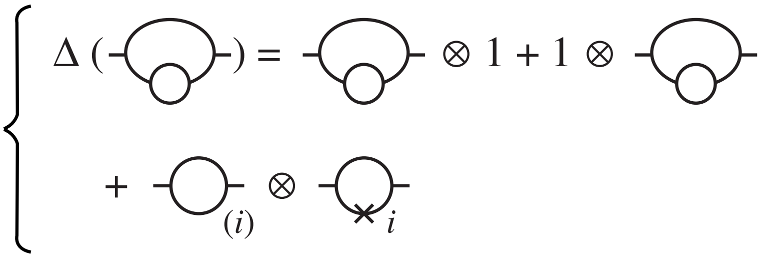
It is useful to see the Hopf algebra of Definition 1.26 in the simplest concrete example. As we did previously, we often use as an example the theory $\mathcal{T} = \phi^3_6$, which has the Lagrangian density (1.137) in dimension $D = 6$. This example is not very physical, both because of the dimension and because the potential does not have a stable equilibrium. However, it is a very convenient example, because it is sufficiently easy (a scalar theory with a very simple potential) and at the same time sufficiently generic from the point of view of renormalization (it is not super-renormalizable).

In this case, the conditions on the class of subgraphs $\mathcal{V}(\Gamma)$ of Definition 1.14 can be rephrased as follows. A subgraph $\gamma \subset \Gamma$ in $\mathcal{V}(\Gamma)$ is a non-trivial (non-empty as well as its complement) subset $\gamma \subset \Gamma^{(1)}_{\rm int}$ such that the components γ_i of γ are 1PI graphs with two or three external lines, together with a choice of index $\chi(\gamma_i) \in \{0, 1\}$ for each component with two external lines. By Proposition 1.15 the graph Γ/γ is a 1PI graph. The map χ is used to give a label $\in \{0, 1\}$ to the two-point vertices of the contracted graph. Thus the above general discussion agrees with that of [**82**].

The following are examples of an explicit calculation of the coproduct for graphs Γ of the ϕ^3_6 theory, taken from [**82**].







As one can see clearly in the examples above, the coproduct has an interesting property of "linearity on the right", which is expressed more precisely in the following result.

Proposition 1.28. ([**82**]) *Let $\mathcal{H}_1$ be the linear subspace of $\mathcal{H}$ generated by 1 and by the 1PI graphs. Then, for all $\Gamma \in \mathcal{H}_1$, the coproduct satisfies*

$$\Delta(\Gamma) \in \mathcal{H} \otimes \mathcal{H}_1 \,.$$

We have introduced the Hopf algebra $\mathcal{H}(\mathcal{T})$ in Definition 1.26 as a Hopf algebra over $\mathbb{C}$. However, for possible arithmetic applications, it is useful to know that, in fact, the Hopf algebras of Feynman graphs can be defined over $\mathbb{Q}$.

Remark 1.29. *The definition of the coproduct and antipode in Definition* 1.26 *continue to make sense if we consider $\mathcal{H}(\mathcal{T})$ be be the free commutative algebra over $\mathbb{Q}$ generated by the 1PI graphs.*

We now discuss several natural gradings of the Hopf algebra $\mathcal{H}(\mathcal{T})$.

Proposition 1.30. *The Hopf algebra $\mathcal{H}(\mathcal{T})$ is graded by the loop number,* $\deg(\Gamma) = b_1(\Gamma)$, *extended by*

$$\deg(\Gamma_1 \cdots \Gamma_r) = \sum_i \deg(\Gamma_i) \quad \textit{and} \quad \deg(1) = 0. \tag{1.199}$$

It also has a grading by the line number

$$\ell(\Gamma) = \#\Gamma^{(1)}_{\mathrm{int}} \quad \textit{and} \quad \ell(\prod_j \Gamma_j) = \sum_j \ell(\Gamma_j). \tag{1.200}$$

The grading by line number has the property that the graded components of $\mathcal{H}(\mathcal{T})$ are finite-dimensional.

PROOF. For any subgraph $\gamma \in \mathcal{V}(\Gamma)$, one has

$$\ell(\Gamma) = \sum_i \ell(\gamma_i) + \ell(\Gamma/\gamma), \tag{1.201}$$

since the set of internal lines of Γ/γ is the complement of γ in $\Gamma^{(1)}_{\rm int}$ while the internal lines of the components are the disjoint subsets γ_i. There are only finitely many graphs with a given number of internal lines. This gives the required finiteness for the dimensions of the graded components. Let us check that the prescription

$$v(\Gamma) = \#\Gamma^{(0)} - 1 \quad \text{and} \quad v(\prod_j \Gamma_j) = \sum_j v(\Gamma_j) \tag{1.202}$$

defines another possible grading on $\mathcal{H}(\mathcal{T})$. We need to show that

$$v(\Gamma) = \sum_i v(\gamma_i) + v(\Gamma/\gamma). \tag{1.203}$$

The graph Γ/γ is obtained by collapsing each of the connected components γ_i to a single vertex, i.e. by replacing the $v(\tilde{\gamma}_i) + 1$ vertices of γ_i with a single vertex. Thus, (1.203) follows. Since both ℓ and v are gradings so is their difference $\ell - v$, which gives the grading by loop number b_1. $\square$

Notice that, with respect to the grading v by number of vertices (1.202), the Hopf algebra $\mathcal{H}(\mathcal{T})$ is in general not connected, i.e. in general the degree-zero component of $\mathcal{H}(\mathcal{T})$ is not just a copy of the field of scalars. One can see this in the example of the ϕ^4 graph of Figure 22.
With the grading given by the loop number, the Hopf algebra $\mathcal{H}(\mathcal{T})$ is graded connected, namely $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$, with $\mathcal{H}_0 = \mathbb{C}$. However, with respect to this choice of grading, the graded components $\mathcal{H}_n$ are in general not finite-dimensional. In fact, one can insert two-point vertices without changing the loop number.
The graded components $\mathcal{H}_n$ are finite-dimensional for the grading ℓ of Proposition 1.30 associated to the counting of the number of internal lines, as shown above. Moreover, for this grading the Hopf algebra is connected, i.e. $\mathcal{H}_0 = \mathbb{C}$.
Thus, as in Lemma 1.20, we have a corresponding affine group scheme that is a projective limit of linear algebraic groups. The affine group scheme is prounipotent, since $\mathcal{H}$ is positively graded connected.

DEFINITION 1.31. *The group of* diffeographisms $\mathrm{Difg}(\mathcal{T})$ *of a renormalizable theory* $\mathcal{T}$ *is the pro-unipotent affine group scheme of the Hopf algebra* $\mathcal{H}(\mathcal{T})$ *of Feynman graphs.*

The terminology "diffeographisms" is motivated by another result of the CK theory, which we discuss in §6.5 below, according to which $\mathrm{Difg}(\mathcal{T})$ acts on the coupling constants of the theory through formal diffeomorphisms tangent to the identity.

The Lie algebra of $\mathrm{Difg}(\mathcal{T})$ is identified in [**82**] using the Milnor–Moore theorem (Theorem 1.22).

THEOREM 1.32. ([**82**]) *The Lie algebra* $\mathrm{Lie}\,\mathrm{Difg}(\mathcal{T})$ *has a canonical linear basis given by pairs* (Γ, w) *consisting of a 1PI graph* Γ *and an element* $w \in J$ *with degree the number of external lines of* Γ. *The Lie bracket is given by*

$$[(\Gamma, w), (\Gamma', w')] = \sum_{v,\ \iota(v)=w'} (\Gamma \circ_v \Gamma', w) - \sum_{v',\ \iota(v')=w} (\Gamma' \circ_{v'} \Gamma, w') \tag{1.204}$$

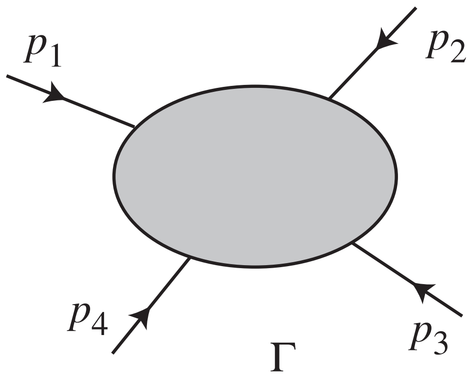


FIGURE 27. The value $U(\Gamma(p_1,\dots,p_N))$ depends on the *incoming* momenta.

where $\Gamma \circ_v \Gamma'$ *denotes the graph obtained by inserting* Γ' *in* Γ *at the vertex* v *of* Γ.

The Lie bracket is uniquely defined from the Milnor–Moore Theorem. In general the notation $\Gamma \circ_v \Gamma'$ is ambiguous since there may be several ways of inserting Γ' in Γ at the vertex v due to the ordering of the external legs, and one has to take into account the corresponding combinatorial factor. We refer to [**276**] for the needed precise discussion of this point.

REMARK 1.33. The Hopf algebra $\mathcal{H}(\mathcal{T})$ simplifies even further when we take the massless case of the ϕ^3 theory in $D = 6$ dimensions (which is consistent in dimension 6). In that case the monomials in the Lagrangian are distinguished by their degrees and one can drop the w as well as the maps ι and χ of the above discussion. Moreover there is no need to introduce the new vertices $\times_{(i)}$ since the only relevant one, namely $\times_{(1)}$, brings in a k^2 which yields the same result as if one eliminates the new vertex altogether (the resulting loss of a propagator is compensated for exactly by the removal of the k^2 vertex). The grading by loop number is now finite in each degree and the counting ℓ of the number of internal lines is no longer a grading.

6.3. The Hopf algebra of Feynman graphs: full structure.

So far, we have not been taking into account the fact that Feynman graphs have an external structure, $\Gamma(p_1,\dots,p_N)$ (cf. Figure 27). We now describe the full Hopf algebra of Feynman graphs, following [**82**].

We need to input the external data in the structure of the algebra. For a Feynman graph Γ of the theory with N external legs, we define the set

$$E_\Gamma := \Big\{(p_i)_{i=1,\dots,N}\ ;\ \sum p_i = 0\Big\} \tag{1.205}$$

of possible external momenta subject to the conservation law. Consider then the space of smooth functions $C^\infty(E_\Gamma)$ and let $C_c^{-\infty}(E_\Gamma)$ denote the space of distributions dual to $C^\infty(E_\Gamma)$.

DEFINITION 1.34. *An external structure on a Feynman graph* Γ *with* N *external legs is a distribution* $\sigma \in C_c^{-\infty}(E_\Gamma)$.

It is often sufficient to consider distributions $\sigma : C^\infty(E_\Gamma) \to \mathbb{C}$ that are δ-functions. However, this is in general not enough. In fact, one should keep in mind the non-linear dependence of $U(\Gamma(p_1,\ldots,p_N))$ on the external momenta.

We choose a set of distributions σ_ι, indexed by the elements $\iota \in J$, so that the σ_ι form a dual basis to the basis of monomials in the full Lagrangian, i.e. they fulfill the analog of (1.166).

We then consider the disjoint union $E := \cup_\Gamma E_\Gamma$, over 1PI graphs Γ and the corresponding direct sum of the spaces of distributions $C_c^{-\infty}(E_\Gamma)$, which we denote by

$$C_c^{-\infty}(E) = \oplus_\Gamma C_c^{-\infty}(E_\Gamma).$$

We obtain the following description of the full Hopf algebra of Feynman graphs.

Definition 1.35. *For a given renormalizable quantum field theory $\mathcal{T}$, the* full Hopf algebra of Feynman graphs *$\tilde{\mathcal{H}}(\mathcal{T})$ is the symmetric algebra on the linear space of distributions $C_c^{-\infty}(E)$,*

$$\tilde{\mathcal{H}}(\mathcal{T}) = \mathrm{Sym}(C_c^{-\infty}(E))\,. \tag{1.206}$$

The coproduct on $\tilde{\mathcal{H}}(\mathcal{T})$ is defined on generators (Γ,σ), with Γ a 1PI graph of $\mathcal{T}$ and $\sigma \in C_c^{-\infty}(E_\Gamma)$, by the formula

$$\Delta(\Gamma,\sigma) = (\Gamma,\sigma)\otimes 1 + 1\otimes(\Gamma,\sigma) + \sum_{\gamma\in\mathcal{V}(\Gamma)} (\gamma,\sigma_{\chi(\gamma)})\otimes(\Gamma/\gamma,\sigma). \tag{1.207}$$

In the formula (1.207), the subgraphs γ belong to the class $\mathcal{V}(\Gamma)$ of Definition 1.14. The notation $(\gamma,\sigma_{\chi(\gamma)})$ is a shorthand for

$$\prod (\tilde{\gamma}_i, \sigma_{\chi(\gamma_i)})$$

where the product is over the components of γ and χ is the map from the set of components of γ to J of Definition 1.14. The contracted graph Γ/γ has the same external lines as the graph Γ so that one can endow it with the external structure σ.

The proof that Definition 1.35 indeed gives a Hopf algebra, cf. [**82**], is a direct extension of the above proof of Theorem 1.27.

The full Hopf algebra $\tilde{\mathcal{H}}(\mathcal{T})$ also has an associated affine group scheme $\widetilde{\mathrm{Difg}}(\mathcal{T})$. The relation of $\widetilde{\mathrm{Difg}}(\mathcal{T})$ to the group of diffeographisms $\mathrm{Difg}(\mathcal{T})$ of the discrete Hopf algebra $\mathcal{H}(\mathcal{T})$ is through a semidirect product

$$\widetilde{\mathrm{Difg}}(\mathcal{T}) = \mathrm{Difg}_{\mathrm{ab}}(\mathcal{T}) \rtimes \mathrm{Difg}(\mathcal{T}), \tag{1.208}$$

where $\mathrm{Difg}_{\mathrm{ab}}(\mathcal{T})$ is abelian. This can be seen at the level of the Lie algebra.

Consider the linear space

$$\mathcal{L} = C^\infty(E), \tag{1.209}$$

with $E := \cup_\Gamma E_\Gamma$, the disjoint union over 1PI graphs Γ. We denote an element of $\mathcal{L}$ as a family $L = (f^\Gamma)$ with $f^\Gamma \in C^\infty(E_\Gamma)$, indexed by 1PI graphs Γ. An element L in $\mathcal{L}$ determines a linear form Z_L on the Hopf algebra $\tilde{\mathcal{H}}(\mathcal{T})$, which is non-zero only for elements of

$$C_c^{-\infty}(E) \subset \mathrm{Sym}(C_c^{-\infty}(E))$$

and pairs as follows with the generators (Γ, σ),

$$\langle Z_L, (\Gamma, \sigma)\rangle := \sigma(f^\Gamma), \tag{1.210}$$

where f^Γ is the component in $C^\infty(E_\Gamma)$ of L. The form $Z_L : \tilde{\mathcal{H}}(\mathcal{T}) \to \mathbb{C}$ satisfies the condition (1.173)

$$Z_L(XY) = Z_L(X)\varepsilon(Y) + \varepsilon(X)Z_L(Y), \quad \forall X, Y \in \tilde{\mathcal{H}}(\mathcal{T}),$$

hence it defines an element in $\operatorname{Lie}\widetilde{\operatorname{Difg}}(\mathcal{T})$. The bracket then corresponds to the commutator $[Z_{L_1}, Z_{L_2}] = Z_{L_1} * Z_{L_2} - Z_{L_2} * Z_{L_1}$, where the product is obtained by transposing the coproduct of $\tilde{\mathcal{H}}(\mathcal{T})$,

$$\langle Z_{L_1} * Z_{L_2}, (\Gamma, \sigma)\rangle = \langle Z_{L_1} \otimes Z_{L_2}, \Delta(\Gamma, \sigma)\rangle.$$

When we evaluate $[Z_{L_1}, Z_{L_2}]$ on (Γ, σ) the only terms of (1.207) which contribute are those coming from subgraphs $\gamma \in \mathcal{V}(\Gamma)$ which have only one component. For such subgraphs $\gamma \in \mathcal{V}_c(\Gamma)$ the map χ is just an element $w \in J$ whose degree is the number of external lines of $\tilde{\gamma}$. One gets

$$\langle [Z_{L_1}, Z_{L_2}], (\Gamma, \sigma)\rangle = \sum_{\gamma \in \mathcal{V}_c(\Gamma),\, w} \sigma_w(f_1^\gamma)\, \sigma(f_2^{\Gamma/\gamma}) - \sigma_w(f_2^\gamma)\, \sigma(f_1^{\Gamma/\gamma})$$

One then has a description of the Lie algebra $\operatorname{Lie}\widetilde{\operatorname{Difg}}(\mathcal{T})$, identified with the Lie algebra of primitive elements in the dual Hopf algebra $\tilde{\mathcal{H}}(\mathcal{T})^\vee$.

Proposition 1.36. *The Lie algebra* $\operatorname{Lie}\widetilde{\operatorname{Difg}}(\mathcal{T})$ *is the linear space* $\mathcal{L}$ *of* (1.209) *endowed with the Lie bracket*

$$[L_1, L_2]^\Gamma = \sum_{\gamma \in \mathcal{V}_c(\Gamma),\, w} \sigma_w(f_1^\gamma)\, f_2^{\Gamma/\gamma} - \sigma_w(f_2^\gamma)\, f_1^{\Gamma/\gamma} \tag{1.211}$$

where $L_j = (f_j^\Gamma)$.

The Lie algebra $\operatorname{Lie}\widetilde{\operatorname{Difg}}(\mathcal{T})$ is the semidirect product of an abelian Lie algebra $\mathcal{L}_{\mathrm{ab}}$ with the Lie algebra $\operatorname{Lie}\operatorname{Difg}(\mathcal{T})$. Elements $L = (f^\Gamma)$ of the abelian subalgebra $\mathcal{L}_{\mathrm{ab}}$ are characterized by the condition

$$L = (f^\Gamma) \in \mathcal{L}_{\mathrm{ab}} \Leftrightarrow \sigma_w(f^\gamma) = 0\,, \quad \forall w \in J \tag{1.212}$$

for any graph γ whose number of external lines is the degree of $w \in J$. We refer the reader to [**82**] for more details.

6.4. BPHZ as a Birkhoff factorization.

We now come to the main result of the CK theory of perturbative renormalization. We begin by describing the Birkhoff factorization of loops.

Definition 1.37. *Let* $C = \partial\Delta$ *denote a circle around the point* $z = 0 \in \mathbb{C}$, *with* Δ *a small disk centered at* $z = 0$. *We denote by* $C_\pm$ *the two components of* $\mathbb{P}^1(\mathbb{C}) \setminus C$, *with* $0 \in C_+$ *and* $\infty \in C_-$. *Let* $G(\mathbb{C})$ *be a connected complex Lie group. Given a smooth loop* $\gamma : C \to G(\mathbb{C})$, *we say that* γ *admits a Birkhoff factorization if it can be written as a product*

$$\gamma(z) = \gamma_-(z)^{-1}\, \gamma_+(z), \quad \forall z \in C\,, \tag{1.213}$$

where the $\gamma_\pm$ *are boundary values of holomorphic functions* $\gamma_\pm : C_\pm \to G(\mathbb{C})$, *with* $\gamma_-(\infty) = 1$.

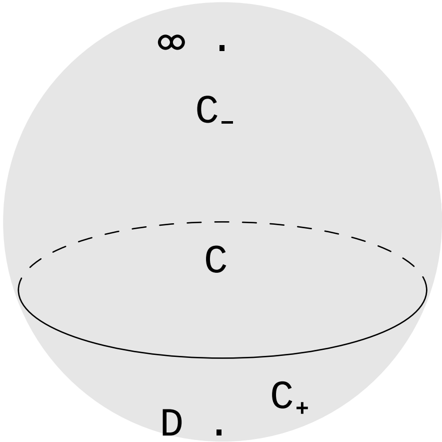


FIGURE 28. Birkhoff decomposition

In general, not all loops admit a Birkhoff factorization. This can be seen in the classical case of loops in $\mathrm{GL}_n(\mathbb{C})$. In this case, the factorization problem is closely related to the classification of holomorphic bundles on $\mathbb{P}^1(\mathbb{C})$. In fact, a loop $\gamma : C \to GL_n(\mathbb{C})$ has in general a factorization of the form

$$\gamma(z) = \gamma_-(z)^{-1}\, \lambda(z)\, \gamma_+(z), \tag{1.214}$$

where $\gamma_\pm$ are boundary values of holomorphic maps $\gamma_\pm : C_\pm \to GL_n(\mathbb{C})$, $\gamma_-(\infty) = 1$, while the middle term λ is a homomorphism of S^1 to the subgroup of diagonal matrices in $\mathrm{GL}_n(\mathbb{C})$,

$$\lambda(z) = \begin{pmatrix} z^{k_1} & & & \\ & z^{k_2} & & \\ & & \ddots & \\ & & & z^{k_n} \end{pmatrix}, \tag{1.215}$$

with integers k_i as exponents. The holomorphic maps $\gamma_\pm$ give the local frames for a trivialization of the restrictions to the open sets $C_\pm$ of a holomorphic vector bundle E on $\mathbb{P}^1(\mathbb{C})$. The bundle E is obtained by gluing these trivializations using the transition function $\lambda : C \to \mathrm{GL}_n(\mathbb{C})$. This recovers in terms of factorizations of loops (cf. [**247**]) the Grothendieck decomposition [**153**] of E as a sum of line bundles

$$E = L_1 \oplus \cdots \oplus L_n\,, \tag{1.216}$$

with Chern classes $c_1(L_i) = k_i$. Up to a permutation of the indices this decomposition is unique and the integers k_i give a complete invariant of the holomorphic vector bundle E.

We are interested in the case when Δ is an infinitesimal disk around $z = 0$ and C an infinitesimal loop. In this case, we can translate the formulation of the Birkhoff factorization (1.213) into the language of affine group schemes.

We assume that the Lie group $G(\mathbb{C})$ is the set of complex points of the affine group scheme G of a commutative Hopf algebra $\mathcal{H}$ over $\mathbb{C}$. Namely, we have $G(\mathbb{C}) = \mathrm{Hom}_{\mathcal{A}_\mathbb{C}}(\mathcal{H}, \mathbb{C})$. We denote by

$$K = \mathbb{C}(\{z\}) = \mathbb{C}\{z\}[z^{-1}] \tag{1.217}$$

the field of convergent Laurent series, that is, germs of meromorphic functions at the origin. We regard K as a unital commutative $\mathbb{C}$-algebra, and consider the K-points of G,

$$G(K) = \mathrm{Hom}_{\mathcal{A}_\mathbb{C}}(\mathcal{H}, K). \tag{1.218}$$

We can think of elements in $G(K)$ as describing loops $\gamma(z)$ on an infinitesimal circle (or an infinitesimal punctured disk Δ^*) around $z = 0$. Similarly, consider

$$\mathcal{O} = \mathbb{C}\{z\}, \tag{1.219}$$

the ring of convergent power series, or germs of holomorphic functions at $z = 0$. Again, we can consider the corresponding group

$$G(\mathcal{O}) = \mathrm{Hom}_{\mathcal{A}_\mathbb{C}}(\mathcal{H}, \mathcal{O}), \tag{1.220}$$

which describes loops $\gamma(z)$ that extend holomorphically to $z = 0$. Finally, we consider the ring

$$\mathcal{Q} = z^{-1}\mathbb{C}[z^{-1}], \quad \text{and the unital ring} \quad \tilde{\mathcal{Q}} = \mathbb{C}[z^{-1}], \tag{1.221}$$

and the group

$$G(\tilde{\mathcal{Q}}) = \mathrm{Hom}_{\mathcal{A}_\mathbb{C}}(\mathcal{H}, \tilde{\mathcal{Q}}). \tag{1.222}$$

We can impose a normalization condition corresponding to $\gamma_-(\infty) = 1$ for an element $\phi \in G(\tilde{\mathcal{Q}})$, by requiring that $\varepsilon_- \circ \phi = \varepsilon$, where ε_- is the augmentation in the ring $\tilde{\mathcal{Q}}$ and ε is the counit of the Hopf algebra $\mathcal{H}$.

In terms of these data, we restate the Birkhoff factorization problem of Definition 1.37 in the following form.

Definition 1.38. *An element $\phi \in G(K)$ admits a Birkhoff factorization if it can be written as a product*

$$\phi = (\phi_- \circ S) * \phi_+ \tag{1.223}$$

with $\phi_+ \in G(\mathcal{O})$ and $\phi_- \in G(\tilde{\mathcal{Q}})$ satisfying $\varepsilon_- \circ \phi_- = \varepsilon$, and with S the antipode of $\mathcal{H}$.

The following result of [**82**] shows that, in the case of interest for renormalization, i.e. when G is the pro-unipotent affine group scheme of a graded connected commutative Hopf algebra $\mathcal{H}$, all loops admit a Birkhoff factorization. In fact, this is not only an existence result, but it provides an explicit recursive formula for the factorization.

Theorem 1.39. ([**82**]) *Let $\mathcal{H}$ be a positively graded connected commutative Hopf algebra, $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$ with $\mathcal{H}_0 = \mathbb{C}$. Then every element $\phi \in G(K) = \mathrm{Hom}_{\mathcal{A}_\mathbb{C}}(\mathcal{H}, K)$ admits a unique Birkhoff factorization, as in* (1.223)*. The factorization is given recursively by the identities*

$$\phi_-(X) = -T\left(\phi(X) + \sum \phi_-(X')\phi(X'')\right), \tag{1.224}$$

where T is the projection on the pole part of the Laurent series, and

$$\phi_+(X) = \phi(X) + \phi_-(X) + \sum \phi_-(X')\phi(X''), \tag{1.225}$$

where, for $X \in \mathcal{H}_n$, the X' and X'' are the lower degree terms that appear in the coproduct

$$\Delta(X) = X \otimes 1 + 1 \otimes X + \sum X' \otimes X''. \tag{1.226}$$

PROOF. We reproduce here the proof from [**82**]. The augmentation ε vanishes on $\mathcal{H}_n$ for $n > 0$, so that the augmentation ideal of $\mathcal{H}$, namely $\mathcal{H}^> = \operatorname{Ker}\varepsilon$, is the direct sum of the $\mathcal{H}_n$ for $n > 0$. For $X \in \mathcal{H}_n$ the coproduct $\Delta(X)$ can be written as a linear combination of tensor products $X' \otimes X''$ of homogeneous elements. The equalities

$$(\mathrm{Id} \otimes \varepsilon) \circ \Delta = \mathrm{Id}\,, \quad (\varepsilon \otimes \mathrm{Id}) \circ \Delta = \mathrm{Id}$$

show that $\Delta(X)$ is of the form (1.226).
One needs to show that the map $\phi_- : \mathcal{H} \to K$ defined by (1.224) is indeed a homomorphism. The main point is to show, by induction, that it is multiplicative. For all X, Y of positive degrees $\deg X > 0$, $\deg Y > 0$ one has

$$\begin{aligned} \Delta(XY) = XY \otimes 1 + 1 \otimes XY + X \otimes Y + Y \otimes X + XY' \otimes Y'' + \\ Y' \otimes XY'' + X'Y \otimes X'' + X' \otimes X''Y + X'Y' \otimes X''Y''\,. \end{aligned} \tag{1.227}$$

Thus, we obtain the following expression for $\phi_-(XY)$:

$$\begin{aligned} \phi_-(XY) = -T(\phi(XY)) - T(\phi_-(X)\,\phi(Y) + \phi_-(Y)\,\phi(X) + \\ \phi_-(XY')\,\phi(Y'') + \phi_-(Y')\,\phi(XY'') + \phi_-(X'Y)\,\phi(X'') \\ + \phi_-(X')\,\phi(X''Y) + \phi_-(X'Y')\,\phi(X''Y''))\,. \end{aligned} \tag{1.228}$$

Now we use the fact that ϕ is a homomorphism and the induction hypothesis that ϕ_- is multipicative, $\phi_-(AB) = \phi_-(A)\,\phi_-(B)$, for $\deg A + \deg B < \deg X + \deg Y$. (Notice that this is automatic if $\deg A = 0$ or $\deg B = 0$ since the algebra is connected and $\phi_-(1) = 1$ so that there is no problem in starting the induction.) We then rewrite (1.228) as

$$\begin{aligned} \phi_-(XY) = \quad & -T(\phi(X)\,\phi(Y) + \phi_-(X)\,\phi(Y) \\ & +\phi_-(Y)\,\phi(X) + \phi_-(X)\,\phi_-(Y')\,\phi(Y'') \\ & +\phi_-(Y')\,\phi(X)\,\phi(Y'') + \phi_-(X')\,\phi_-(Y)\,\phi(X'') \\ & +\phi_-(X')\,\phi(X'')\,\phi(Y) \\ & +\phi_-(X')\,\phi_-(Y')\,\phi(X'')\,\phi(Y''))\,. \end{aligned} \tag{1.229}$$

We need to compare this with $\phi_-(X)\phi_-(Y)$. We can compute the latter using the fact that the operator of projection onto the polar part satisfies the relation

$$T(f)\,T(h) = -T(fh) + T(T(f)\,h) + T(f\,T(h))\,. \tag{1.230}$$

By applying (1.230) to

$$f = \phi(X) + \phi_-(X')\,\phi(X'') \qquad \text{and} \qquad h = \phi(Y) + \phi_-(Y')\,\phi(Y''),$$

we obtain

$$
\begin{aligned}
&\phi_-(X)\,\phi_-(Y) = \\
&-T\Big(\big(\phi(X)+\phi_-(X')\,\phi(X'')\big)\,\big(\phi(Y)\;+\phi_-(Y')\,\phi(Y'')\big)\Big) \\
&+T\Big(T\big(\phi(X)+\phi_-(X')\,\phi(X'')\big)\,\big(\phi(Y)\;+\phi_-(Y')\,\phi(Y'')\big)\Big) \\
&+T\Big(\big(\phi(X)+\phi_-(X')\,\phi(X'')\big)\,T\big(\phi(Y)+\phi_-(Y')\,\phi(Y'')\big)\Big).
\end{aligned}
\tag{1.231}
$$

Using the fact that $T(f) = -\phi_-(X)$, $T(h) = -\phi_-(Y)$, we can then rewrite (1.231) in the form

$$
\begin{aligned}
&\phi_-(X)\,\phi_-(Y) = \\
&-T\Big(\phi(X)\,\phi(Y) + \phi_-(X')\,\phi(X'')\,\phi(Y) \\
&+\phi(X)\,\phi_-(Y')\,\phi(Y'')\;\;+\phi_-(X')\,\phi(X'')\,\phi_-(Y')\,\phi(Y'')\Big) \\
&-T\Big(\phi_-(X)(\phi(Y)+\phi_-(Y')\,\phi(Y'')\Big) \\
&-T\Big(\big(\phi(X)+\phi_-(X')\,\phi(X'')\big)\,\phi_-(Y)\Big)\,.
\end{aligned}
\tag{1.232}
$$

Finally, we compare (1.229) with (1.232). Both expressions contain eight terms of the form $-T(a)$ and one can check that they match in pairs. This proves that ϕ_- defined by the recursive formula (1.224) is indeed an algebra homomorphism. It is then clear that the ϕ_+ defined by (1.225) is also an algebra homomorphism, since both ϕ and ϕ_- are. It is clear by construction that $\phi_- \in G(\tilde{\mathcal{Q}})$ since it is a polar part, and that $\phi_+ \in G(\mathcal{O})$. It is also easy to check that one has $\phi_+ = \phi_- * \phi$, since one can see the right-hand side of (1.225) as the pairing $\langle \phi_- \otimes \phi, \Delta(X)\rangle = \phi_- * \phi\,(X)$. $\square$

One can then apply the result of Theorem 1.39 to the specific case of the Hopf algebra of Feynman graphs. One finds that the formulae (1.224) and (1.225) for the recursive Birkhoff factorization of loops in the group of diffeographisms are exactly the recursive formulae of the BPHZ renormalization procedure.

THEOREM 1.40. ([**82**]) *Let $\mathcal{T}$ be a renormalizable theory and $\tilde{\mathcal{H}}(\mathcal{T})$ the Hopf algebra of Feynman graphs, with the grading by loop number. Then the formulae* (1.224) *and* (1.225) *for the Birkhoff factorization of loops $\phi \in \widetilde{\mathrm{Difg}}(\mathcal{T})(K) = \mathrm{Hom}_{\mathcal{A}_{\mathbb{C}}}(\tilde{\mathcal{H}}(\mathcal{T}), K)$ are given, respectively, by the formulae* (1.163) *and* (1.165) *of the BPHZ perturbative renormalization.*

PROOF. Consider the data $U^z(\Gamma(p_1,\ldots,p_N))$ of the unrenormalized values of Feynman graphs, regularized by applying DimReg. We can view U as a homomorphism $U : \tilde{\mathcal{H}}(\mathcal{T}) \to K$, which assigns to a generator (Γ, σ) of $\tilde{\mathcal{H}}(\mathcal{T})$ the Laurent series

$$
h(z) = \langle \sigma, U^z(\Gamma(p_1,\ldots,p_N))\rangle. \tag{1.233}
$$

Here we view $U^z(\Gamma(p_1,\dots,p_N)) = f_z(p_1,\dots,p_N)$ as a family of smooth functions $f_z \in C^\infty(E_\Gamma)$, with $z \in \Delta^*$ varying in an infinitesimal punctured disk around $z = 0$. Thus, we can pair the distribution σ with f_z and obtain $h(z) = \langle \sigma, f_z \rangle \in K = \mathbb{C}(\{z\})$. It is then clear that, upon setting $\phi_- = C$ and $\phi_+ = R$ one can identify the formulae (1.163) and (1.165) with (1.224) and (1.225). □

Notice that it is fine here to consider the grading by loop number because for the inductive procedure connectedness of the Hopf algebra suffices and one does not need it to be finite-dimensional in each degree.

We can express the result of Theorem 1.40 equivalently in terms of loops γ. We say that the data $U^z(\Gamma(p_1,\dots,p_N))$ define a loop $\gamma(z)$ in the pro-unipotent Lie group of complex points of $\widetilde{\mathrm{Difg}}(\mathcal{T})$. This loop has a Birkhoff factorization

$$\gamma(z) = \gamma_-(z)^{-1}\gamma_+(z),$$

where the "negative piece" $\gamma_-(z)$ gives the counterterms and the "positive part" $\gamma_+(z)$ gives the renormalized value as the evaluation $\gamma_+(0)$.

We now restrict attention to the discrete Hopf algebra $\mathcal{H}(\mathcal{T})$ for the simple reason that the non-trivial counterterms in the full Hopf algebra already come from the Birkhoff factorization of the restriction to $\mathcal{H}(\mathcal{T})$. In fact, notice that terms of the form

$$(\gamma, \sigma_{\chi(\gamma)}) \otimes (\Gamma/\gamma, \sigma)$$

in the coproduct formula of the full Hopf algebra $\tilde{\mathcal{H}}(\mathcal{T})$ are identified with elements in $\mathcal{H}(\mathcal{T})$ through the choice of the dual basis σ_ι. Thus, if one knows the value of the restriction of $\gamma_-(z)$ to $\mathcal{H}(\mathcal{T})$, the renormalized value of an arbitrary graph does not require the computation of any new counterterm. It is obtained, after preparation by the known counterterms (from $\mathcal{H}(\mathcal{T})$), by a simple subtraction of the pole part (which is non-zero only if the degree of divergence is positive).

THEOREM 1.41. *Let $\mathcal{T}$ be a renormalizable theory and $\mathcal{H}(\mathcal{T})$ the discrete Hopf algebra of Feynman graphs, with its affine group scheme* $\mathrm{Difg}(\mathcal{T})$. *Then the formulae* (1.224) *and* (1.225) *for the Birkhoff factorization of loops applied to the homomorphism U give respectively the counterterms and the renormalized value for the theory $\mathcal{T}'$ with new interaction vertices associated to monomials of the non-interacting part of the Lagrangian of $\mathcal{T}$.*

PROOF. We let as above J be the set of all monomials in the Lagrangian of $\mathcal{T}$ and I the set of interaction monomials. For the graphs $\Gamma \in \mathrm{Graph}\,\mathcal{T}$ such that $\iota(\Gamma) \subset I$ the statement follows from Proposition 1.40.

The formulae (1.224) and (1.225) depend upon the definition of subgraphs of the theory, through the fact that they involve the coproduct in the Hopf algebra. We need to show that our definition of subgraphs has the right properties so that (1.224) and (1.225) give the renormalization of the extended Feynman graphs of the theory (as in Definition 1.11) with $\iota(\Gamma^{(0)}) \subset J$, viewing these graphs as Feynman graphs of the new theory $\mathcal{T}'$.

The main point is that the new theory is still renormalizable and does not require any new type of counterterms in order to be renormalized. This is proved using dimensional analysis. When all the coupling constants are positive one gets a finite number of possible counterterms (cf. [**62**] §3.3.3).

For instance, in the theory ϕ^3 in dimension D the field ϕ has physical dimension $D/2-1$ and the coupling constant g has physical dimension $3-D/2$ (cf. §2.2 above). As long as $D \leq 6$, the theory is renormalizable and it remains such with the new vertices $\overset{0}{-\!\!\times\!\!-}$ and $\overset{1}{-\!\!\times\!\!-}$. Moreover, when we renormalize a graph of that new theory, we do not introduce new counterterms. The three types of counterterms needed are already present from $\mathcal{T}$.

Now start from $\Gamma \in \mathrm{Graph}(\mathcal{T})$ and view it as a graph of the new theory $\mathcal{T}'$. Then, in order to renormalize it, we first prepare it by adding the counterterms corresponding to the subdivergences. We know that the only subgraphs responsible for subdivergences of Γ have at most three external legs (in general they belong to the set J), hence they correspond to subgraphs in the sense of Definition 1.14 for the theory $\mathcal{T}$. According to this definition, we use already all the elements of J as labels for the connected components of the subgraphs. This means that, when we collect the terms, we get the prepared graph as if we were applying BPHZ for the new theory $\mathcal{T}'$. One then proceeds by induction to show that in (1.224) we get the counterterm for Γ as if we were applying BPHZ for the new theory $\mathcal{T}'$. The same holds for the renormalized value. It is crucial, in doing this, that we did not need to add new vertices since the propagator of the new theory remains the same. $\square$

6.5. Diffeographisms and diffeomorphisms.

Another important result of the CK theory, which we discuss in this section, shows that the group of diffeographisms of a renormalizable theory $\mathcal{T}$ maps to the group of formal diffeomorphisms tangent to the identity of the space of coupling constants of the theory.

More precisely, for a renormalizable theory $\mathcal{T}$, consider the complex vector space V with a basis labeled by the coupling constants. We let $\mathrm{Diff}(V)$ be the group of formal diffeomorphisms of V tangent to the identity at $0 \in V$ and $\mathcal{H}_{\mathrm{diff}}(V)$ be its Hopf algebra. We work over the ground field $\mathbb{C}$. There is a Hopf algebra homomorphism $\Phi : \mathcal{H}_{\mathrm{diff}}(V) \to \mathcal{H}(\mathcal{T})$ and a dual group homomorphism $\mathrm{Difg}(\mathcal{T}) \to \mathrm{Diff}(V)$. This map to formal diffeomorphisms explains the terminology "diffeographisms" for the group scheme $\mathrm{Difg}(\mathcal{T})$.

The map Φ is constructed by assigning to the coefficients of the expansion of formal diffeomorphisms the coefficients in $\mathcal{H}(\mathcal{T})$ of the expansion of the effective coupling constants of the theory as formal power series in the bare coupling constants. It is not immediate to show that Φ is indeed a Hopf algebra homomorphism ([**83**]).

We state this result more precisely in the simplest case where $\mathcal{T}$ is the massless ϕ^3_6, where there is only one coupling constant, and one can write a completely explicit formula for the map Φ.

Proposition 1.42. ([**83**]) *Let $\mathcal{T}$ be the massless ϕ^3_6 theory and let $\mathcal{H}_{\mathrm{diff}}(\mathbb{C})$ be the Hopf algebra of formal diffeomorphisms of $\mathbb{C}$ tangent to the identity. Then the following holds.*

(1) *The effective coupling constant has a power series expansion of the form*

$$g_{\text{eff}} = \frac{\left(g + \displaystyle\sum g^{2\ell+1} \frac{\Gamma}{\sigma(\Gamma)} \right)}{\left(1 - \displaystyle\sum g^{2\ell} \frac{\Gamma}{\sigma(\Gamma)} \right)^{3/2}}, \tag{1.234}$$

where the coefficients are elements in the Hopf algebra $\mathcal{H}(\mathcal{T})$, with $\sigma(\Gamma)$ the symmetry factor of the graph as in (1.98).

(2) *The expansion* (1.234) *induces a Hopf algebra homomorphism $\Phi : \mathcal{H}_{\text{diff}}(\mathbb{C}) \to \mathcal{H}(\phi_6^3)$ with $\Phi(a_n) = \alpha_n$.*

The expansion (1.234) is obtained by considering the effect of adjusting the coupling constants in the Lagrangian

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 - \frac{g}{6}\phi^3$$

of the massless ϕ_6^3 theory by

$$\frac{1}{2}(\partial\phi)^2(1 - \delta Z) - \frac{g + \delta g}{6}\phi^3.$$

The correction $\frac{1}{2}(\partial\phi)^2(1 - \delta Z)$ can be absorbed in a rescaling $\phi \mapsto \phi(1 - \delta Z)^{1/2}$ so that one obtains

$$\frac{1}{2}(\partial\phi)^2 - \frac{1}{6}(g + \delta g)(1 - \delta Z)^{-3/2}\,\phi^3.$$

The counterterms responsible for the term δZ come from all the graphs with two external legs, while the counterterms involved in the term δg come from graphs with three external legs (cf. the BPHZ formulae (1.162), (1.163), (1.165)), hence the explicit form of the expansion (1.234).

The Hopf algebra $\mathcal{H}_{\text{diff}}(\mathbb{C})$ of formal diffeomorphisms of $\mathbb{C}$ tangent to the identity has generators the coordinates a_n of

$$\varphi(x) = x + \sum_{n \geq 2} a_n(\varphi) x^n,$$

for φ a formal diffeomorphism satisfying $\varphi(0) = 0$ and $\varphi'(0) = \text{id}$. The coproduct $\Delta(a_n)$ is given by

$$\langle \Delta(a_n), \varphi_1 \otimes \varphi_2 \rangle = a_n(\varphi_2 \circ \varphi_1).$$

One obtains an algebra homomorphism $\Phi : \mathcal{H}_{\text{diff}}(\mathbb{C}) \to \mathcal{H}(\phi_6^3)$ by setting $\Phi(a_n) = \alpha_n$, where the $\alpha_n \in \mathcal{H}(\phi_6^3)$ are the coefficients of the series $g_{\text{eff}} = g + \sum_{n \geq 2} \alpha_n, g^n$, which satisfy $\alpha_{2n} = 0$. We refer the reader to [**83**] §4 for the proof that the algebra homomorphism Φ is also comultiplicative.

The fact that there is a group representation $\text{Difg}(\mathcal{T}) \to \text{Diff}(V)$ means that it is possible to lift at the level of diffeographisms various aspects of the theory of renormalization. We will see in §6.6 below that the beta function and the action of the renormalization group can be formulated naturally at the level of the group scheme of diffeographisms.

This result was generalized from the case of massless ϕ^3_6 theory to other theories by Cartier and by Krajewski.

6.6. The renormalization group.

Recall how we remarked in §4 that, in fact, the regularized unrenormalized values

$$U^z(\Gamma(p_1, \ldots, p_N))$$

depend on an additional mass parameter μ, as in (1.135). This induces a corresponding dependence on the parameter μ of the loop $\gamma_\mu(z)$ that encodes the unrenormalized values as in Proposition 1.40. For a graph Γ with loop number L this dependence is simply given by a multiplicative factor of μ^{zL} when working in dimension $D - z$. In the example of the theory $\mathcal{T} = \phi^3_6$, this dependence on μ is described explicitly in terms of the Feynman rules in [**87**] §2.7. Our convention here agrees with that of [**87**] on the relevant Hopf algebra of graphs with two or three external legs, which are the only ones with non-zero counterterms. We discuss here two general properties of $\gamma_\mu(z)$; see Propositions 1.43 and 1.44 below.

Consider the 1-parameter group of automorphisms

$$\theta_t \in \mathrm{Aut}(\mathrm{Difg}(\mathcal{T})), \quad \forall t \in \mathbb{C} \tag{1.235}$$

implementing the grading by loop number, namely, with infinitesimal generator

$$\frac{d}{dt}\, \theta_t \,|_{t=0} = Y \tag{1.236}$$

given by the grading operator

$$Y(X) = n\, X, \quad \forall X \in \mathcal{H}^\vee_n(\mathcal{T}), \tag{1.237}$$

so that

$$\theta_t(X) = e^{nt}\, X, \quad \forall X \in \mathcal{H}^\vee_n(\mathcal{T})\,,\ t \in \mathbb{C}. \tag{1.238}$$

We let the grading θ_t act by automorphisms of both $\mathcal{H} = \mathcal{H}(\mathcal{T})$ and the dual algebra $\mathcal{H}^\vee$ so that

$$\langle \theta_t(u), x\rangle = \langle u, \theta_t(x)\rangle\,, \quad \forall x \in \mathcal{H}\,,\ \forall u \in \mathcal{H}^\vee\,. \tag{1.239}$$

We obtain the following result about the dependence of $\gamma_\mu(z)$ on the parameter μ.

PROPOSITION 1.43. ([**83**]) *Let $\gamma_\mu(z)$ be the loop in the pro-unipotent Lie group of diffeographisms that encodes the regularized unrenormalized values $U^z_\mu(\Gamma)$. Let θ_t be the 1-parameter family of automorphisms* (1.235)*. Then, for all $t \in \mathbb{R}$ and for all $z \in \Delta^*$, the loop $\gamma_\mu(z)$ satisfies the scaling property*

$$\gamma_{e^t\mu}(z) = \theta_{tz}(\gamma_\mu(z)). \tag{1.240}$$

PROOF. Since both sides of (1.240) are homomorphisms from $\mathcal{H}$ to K it is enough to check the equality on the generators Γ and one can use (1.239). It corresponds to the fact that the dependence on μ in the unrenormalized value of the graph Γ (1.135) is through a power μ^{zL}, with L the loop number $L = b_1(\Gamma)$, which comes from formally replacing the integration variables $d^D k$ by $\mu^z d^{D-z} k$ in the "integration in dimension $D - z$" of (1.132). □

The following result ([**62**], §5.8 and §7.1) will play a very important role in §7 below.

PROPOSITION 1.44. ([**83**]) *Let $\gamma_\mu(z)$ be the loop in the pro-unipotent Lie group of diffeographisms that encodes the regularized unrenormalized values $U^z_\mu(\Gamma)$. Let*

$$\gamma_\mu(z) = \gamma_{\mu^-}(z)^{-1}\gamma_{\mu^+}(z)$$

be the Birkhoff factorization of Theorem 1.39. Then the negative part of the Birkhoff factorization is independent of the mass parameter μ,

$$\frac{\partial}{\partial\mu}\gamma_{\mu^-}(z) = 0\,. \tag{1.241}$$

PROOF. Let us give the proof in some detail. First notice that we are dealing with general graphs with vertices in J, rather than the restricted graphs with vertices in I. By Theorem 1.41, when evaluated on a graph $\Gamma \in \mathrm{Graph}\,\mathcal{T}$, the negative part $\gamma_{\mu^-}(z)$ of the Birkhoff factorization is the same as the counterterms of that graph in the BPHZ procedure of the new theory $\mathcal{T}'$ with the new vertices labeled by monomials in $J \setminus I$. Notice that these new vertices have coupling constants of positive physical dimension (in the sense of §2.2 above), which are independent of μ. Thus, we are reduced to the analysis of the μ-dependence for counterterms in a renormalizable theory (with couplings of physical dimensions ≥ 0). The proof that the counterterms are independent of μ is done in three steps, as follows.

- The counterterms depend in a polynomial manner on the mass parameters of the theory (μ not included).
- Only powers of $\log\mu$ can appear in the counterterms.
- By dimensional analysis the counterterms do not depend on μ.

We refer to [**62**] §5.8.1 for the proof of the first statement. It is based on repeated differentiation with respect to these parameters and an inductive argument assuming that the counterterms for the subdivergences are already polynomial. This behavior is specific to DimReg and MS and would not apply, for instance, in the mass-shell renormalization scheme.

That only powers of $\log\mu$ can appear in the counterterms follows again by an inductive argument since the counterterms are obtained by extracting the pole part of an expression with new dependence in μ of the form μ^{zL} which one expands in powers of z and $\log\mu$.

The argument of a log is necessarily dimensionless since when we apply the minimal subtraction we are subtracting quantities which have the same dimension. Thus the massive parameter μ can only appear in the form of products of $\log(p^2/\mu^2)$ or of $\log(M^2/\mu^2)$, where M is a mass parameter of the theory. However, we know from the discussion above that momenta p or such mass parameters can only appear as polynomials, hence such expressions of logarithmic form are excluded and μ cannot appear in the counterterms. ◻

DEFINITION 1.45. *In the following, let $G(\mathbb{C})$ be the pro-unipotent affine group scheme associated to a positively graded connected commutative Hopf algebra $\mathcal{H}$. Let $L(G(\mathbb{C}),\mu)$ denote the space of $G(\mathbb{C})$-valued loops $\gamma_\mu(z)$, defined on an infinitesimal punctured disk $z \in \Delta^*$ around the origin, and satisfying the scaling condition* (1.240) *and the property* (1.241) *that the negative part of the unique Birkhoff factorization of $\gamma_\mu(z)$ is independent of μ.*

The results discussed in the rest of this section apply to any $\gamma_\mu \in L(G(\mathbb{C}), \mu)$, hence in particular to the case when $G(\mathbb{C})$ is the group of complex points of $\mathrm{Difg}(\mathcal{T})$ and $\gamma_\mu(z)$ encodes the regularized unrenormalized values $U^z_\mu(\Gamma)$ of the theory.

DEFINITION 1.46. *Suppose given $\gamma_\mu(z) \in L(G(\mathbb{C}), \mu)$. The residue is defined as*

$$\mathrm{Res}_{z=0}\,\gamma := -\left(\frac{\partial}{\partial u}\gamma_-\left(\frac{1}{u}\right)\right)_{u=0}. \tag{1.242}$$

The corresponding beta function is then given by

$$\beta := Y\,\mathrm{Res}\,\gamma, \tag{1.243}$$

where Y is the grading operator (1.237).

Since (1.242) depends only on γ_-, the residue is independent of μ by Proposition 1.44. The beta function β of (1.243) is an element in the Lie algebra $\mathrm{Lie}\,G$.

The renormalization group is obtained as a 1-parameter subgroup F_t of $G(\mathbb{C})$ as follows.

PROPOSITION 1.47. ([**83**]) *Suppose given $\gamma_\mu(z) \in L(G(\mathbb{C}), \mu)$. Then the following properties hold.*

(1) *The loop $\gamma_-(z)\,\theta_{tz}(\gamma_-(z)^{-1})$ is regular at $z = 0$.*
(2) *The limit*

$$F_t = \lim_{z\to 0} \gamma_-(z)\,\theta_{tz}(\gamma_-(z)^{-1}) \tag{1.244}$$

defines a 1-parameter subgroup of $G(\mathbb{C})$. Viewed as a homomorphism

$$F_t \in \mathrm{Hom}_{\mathcal{A}_\mathbb{C}}(\mathcal{H}, \mathbb{C}),$$

it has the property that $F_t(X)$ is polynomial in t, for all $X \in \mathcal{H}$.
(3) *The infinitesimal generator of F_t is the beta function* (1.243).
(4) *Let $\gamma_{\mu^+}(z)$ be the positive part of the Birkhoff factorization of $\gamma_\mu(z)$. It satisfies*

$$\gamma_{e^t\mu^+}(0) = F_t\,\gamma_{\mu^+}(0)\,, \quad \forall t \in \mathbb{R}. \tag{1.245}$$

PROOF. (1) Since $\gamma_-(z)$ is the negative part in the Birkhoff factorization of $\gamma_\mu(z)$ and of $\theta_{-tz}(\gamma_\mu(z)) = \gamma_{e^{-t}\mu}(z)$, the elements $\gamma_-(z)\,\gamma_\mu(z)$ and $y(z) := \gamma_-(z)\,\theta_{-tz}(\gamma_\mu(z))$ are regular at $z = 0$. It follows that $\theta_{tz}(y(z)) = \theta_{tz}(\gamma_-(z))\,\gamma_\mu(z)$ is regular at $z = 0$, hence so is the ratio

$$\gamma_-(z)\,\theta_{tz}(\gamma_-(z)^{-1}) = (\gamma_-(z)\,\gamma_\mu(z))(\theta_{tz}(\gamma_-(z))\,\gamma_\mu(z))^{-1}.$$

(2) Thus, for any $t \in \mathbb{R}$, the limit

$$\lim_{z\to 0}\,\langle \gamma_-(z)\,\theta_{tz}(\gamma_-(z)^{-1}), X\rangle \tag{1.246}$$

exists, for any $X \in \mathcal{H}$. Here we view the element $\gamma_-(z)\,\theta_{tz}(\gamma_-(z)^{-1})$ of $G(\mathbb{C})$ as a homomorphism $\mathcal{H} \to \mathbb{C}$.

The 1-parameter family θ_t of automorphisms of G (cf. (1.235)) also acts as automorphisms of the Hopf algebra $\mathcal{H}$, satisfying by definition

$$\langle \theta_t(\gamma), X\rangle = \langle \gamma, \theta_t(X)\rangle\,, \quad \forall X \in \mathcal{H}\,, \quad \forall \gamma \in G(\mathbb{C})\,. \tag{1.247}$$

Thus, for $X \in \mathcal{H}$, we obtain

$$\langle \gamma_-(z)\,\theta_{tz}(\gamma_-(z)^{-1}), X\rangle = \langle \gamma_-(z)^{-1} \otimes \gamma_-(z)^{-1}, (S \otimes \theta_{tz})\,\Delta(X)\rangle\,. \tag{1.248}$$

Upon writing the coproduct as a sum of homogeneous elements

$$\Delta(X) = \sum X_{(1)} \otimes X_{(2)},$$

we obtain (1.248) as a sum of terms of the form

$$\langle \gamma_-(z)^{-1}, S\, X_{(1)} \rangle \, \langle \gamma_-(z)^{-1}, \theta_{tz}\, X_{(2)} \rangle = P_1\left(\frac{1}{z}\right)\, e^{ktz}\, P_2\left(\frac{1}{z}\right), \tag{1.249}$$

for polynomials P_1, P_2. We know, by the existence of the limit (1.246), that the sum of these terms is holomorphic at $z = 0$. Thus, by expanding e^{ktz} in a Taylor series, we obtain that

$$\langle F_t, X\rangle = \lim_{z\to 0} \langle \gamma_-(z)\, \theta_{tz}(\gamma_-(z)^{-1}), X\rangle$$

is a polynomial in t.

To see that F_t is a one-parameter subgroup of $G(\mathbb{C})$, we endow $G(\mathbb{C})$ with the topology of pointwise convergence

$$\gamma_n \to \gamma \quad \text{iff} \quad \langle \gamma_n, X\rangle \to \langle \gamma, X\rangle, \quad \forall X \in \mathcal{H}\,. \tag{1.250}$$

The fact that $G(\mathbb{C})$ is a topological group follows from the equality

$$\langle \gamma\, \gamma', X\rangle = \sum \langle \gamma, X_{(1)}\rangle \, \langle \gamma', X_{(2)}\rangle$$

and similarly for the inverse. Using (1.247) and the definition of F_t we have

$$\lim_{z\to 0} \theta_{sz}(\gamma_-(z)\, \theta_{tz}(\gamma_-(z)^{-1})) = F_t, \tag{1.251}$$

hence we obtain

$$\begin{aligned} F_{s+t} &= \lim_{z\to 0}\ \gamma_-(z)\, \theta_{(s+t)z}\, (\gamma_-(z)^{-1}) \\ &= \lim_{z\to 0}\ \gamma_-(z)\, \theta_{sz}(\gamma_-(z)^{-1})\, \theta_{sz}(\gamma_-(z)\, \theta_{tz}(\gamma_-(z)^{-1})) = F_s\, F_t\,, \end{aligned}$$

so that $F_{s+t} = F_s F_t$, for all $s, t \in \mathbb{R}$.

(3) This will follow from Lemma 1.48, cf. Corollary 1.49 below.

(4) The value $\gamma_\mu^+(0)$ is the regular value of $\gamma_-(z)\, \gamma_\mu(z)$ at $z = 0$. Similarly, $\gamma^+_{e^t\mu}(0)$ is the regular value of $\gamma_-(z)\, \theta_{tz}(\gamma_\mu(z))$, or equivalently of $\theta_{-tz}(\gamma_-(z))\, \gamma_\mu(z)$, at $z = 0$. We know that the ratio satisfies

$$\theta_{-tz}(\gamma_-(z))\, \gamma_-(z)^{-1} \to F_t$$

when $z \to 0$, hence we obtain the result. □

We conclude this section by the following result of [**83**], which will be the starting point for the main topic of the next Section, the Riemann–Hilbert correspondence underlying perturbative renormalization.

LEMMA 1.48. *Suppose given $\gamma_\mu(z) \in L(G(\mathbb{C}), \mu)$ and let $\gamma_-(z)$ be the negative piece of the Birkhoff factorization, which can be written as*

$$\gamma_-(z)^{-1} = 1 + \sum_{n=1}^{\infty} \frac{d_n}{z^n}, \tag{1.252}$$

with coefficients $d_n \in \mathcal{H}^\vee$. Then the coefficients d_n satisfy the relations

$$Y\,(d_{n+1}) = d_n\, \frac{d}{dt} F_t|_{t=0} \quad \forall n \geq 1, \quad \textit{and} \quad Y(d_1) = \frac{d}{dt} F_t|_{t=0}\,. \tag{1.253}$$

PROOF. We first show that, for all $X \in \mathcal{H}$, we have

$$\langle \frac{d}{dt} F_t|_{t=0}, X\rangle = \lim_{z\to 0} z\langle \gamma_-(z)^{-1} \otimes \gamma_-(z)^{-1}, (S\otimes Y)\,\Delta(X)\rangle. \tag{1.254}$$

We know from (1.248) and (1.244) that, for $z \to 0$, we have

$$\langle \gamma_-(z)^{-1} \otimes \gamma_-(z)^{-1}, (S\otimes \theta_{tz})\,\Delta(X)\rangle \to \langle F_t, X\rangle. \tag{1.255}$$

We denote by Ξ the expression on the left-hand side of (1.255). We know by (1.249) that it is a finite sum $\Xi = \sum P_k(z^{-1})\, e^{ktz}$, where the P_k are polynomials. As before, we can replace the exponentials e^{ktz} by their Taylor expansions up to some order N in z and obtain polynomials in t with coefficients that are Laurent series in z. Since we know the expression is regular at $z = 0$ for any value of t, these polynomials have, in fact, coefficients that are regular at $z = 0$, i.e. that are polynomials in z. Thus, the left-hand side of (1.255) is a uniform family of holomorphic functions of t, for t varying in a disk around $t = 0$. The derivative $\partial_t\Xi$ at $t = 0$ converges to $\partial_t\langle F_t, X\rangle|_{t=0}$ when $z \to 0$, so that we have

$$z\,\langle \gamma_-(z)^{-1} \otimes \gamma_-(z)^{-1}, (S\otimes Y)\,\Delta(X)\rangle \to \langle \frac{d}{dt} F_t|_{t=0}, X\rangle. \tag{1.256}$$

Notice then that the function

$$z \mapsto z\,\langle \gamma_-(z)^{-1} \otimes \gamma_-(z)^{-1}, (S\otimes Y)\,\Delta(X)\rangle \tag{1.257}$$

is holomorphic for $z \in \mathbb{C}\backslash\{0\}$ and extends holomorphically to $z = \infty \in \mathbb{P}_1(\mathbb{C})$, since $\gamma_-(\infty) = 1$, so that $Y(\gamma_-(\infty)) = 0$ (see (1.259)). Moreover, by the previous argument, it is also holomorphic at $z = 0$, which implies that it is a constant. This gives, for all $X \in \mathcal{H}$,

$$\langle \gamma_-(z)^{-1} \otimes \gamma_-(z)^{-1}, (S\otimes Y)\,\Delta(X)\rangle = \frac{1}{z}\,\langle \frac{d}{dt} F_t|_{t=0}, X\rangle\,,$$

which is equivalent to

$$\langle \gamma_-(z)\,Y(\gamma_-(z)^{-1}), X\rangle = \frac{1}{z}\,\langle \frac{d}{dt} F_t|_{t=0}, X\rangle.$$

Thus, one obtains the identity

$$Y(\gamma_-(z)^{-1}) = \frac{1}{z}\,\gamma_-(z)^{-1}\,\frac{d}{dt} F_t|_{t=0}\,. \tag{1.258}$$

Using (1.252), one obtains

$$Y(\gamma_-(z)^{-1}) = \sum_{n=1}^{\infty} \frac{Y(d_n)}{z^n} \tag{1.259}$$

and

$$\frac{1}{z}\,\gamma_-(z)^{-1}\,\frac{d}{dt} F_t|_{t=0} = \frac{1}{z}\,\frac{d}{dt} F_t|_{t=0} + \sum_{n=1}^{\infty} \frac{d_n}{z^{n+1}}\,\frac{d}{dt} F_t|_{t=0}.$$

These identities applied to the two sides of (1.258) yield the desired relations (1.253). □

COROLLARY 1.49. *The beta function defined in* (1.243) *of Definition* 1.46 *is the infinitesimal generator of the renormalization group F_t of Proposition* 1.47, *namely*

$$\frac{d}{dt} F_t|_{t=0} = \beta. \tag{1.260}$$

PROOF. It is sufficient to see that in the equation (1.253) we have $Y(d_1) = \frac{d}{dt}F_t|_{t=0}$ and that d_1 is by construction the residue $\mathrm{Res}_{z=0}\gamma$. The minus sign in (1.242) accounts for the use of $\gamma_-(z)$ instead of its inverse $\gamma_-(z)^{-1}$. Thus, by (1.243) we have $Y(d_1) = \beta$, which gives (1.260). □

7. Renormalization and the Riemann–Hilbert correspondence

> *"La parenté de plus en plus manifeste entre le groupe de Grothendieck–Teichmüller d'une part, et le groupe de renormalisation de la Théorie Quantique des Champs n'est sans doute que la première manifestation d'un groupe de symétrie des constantes fondamentales de la physique, une espèce de groupe de Galois cosmique!"*
>
> Pierre Cartier, [**42**]

In this passage, written in 2000, Cartier conjectured the existence of a hidden group of symmetries, closely related in structure to certain arithmetic Galois groups, and acting on the constants of physical theories in a way related to the action of the renormalization group.

In this section, which is based on the result of our work [**87**], [**89**], we verify this conjecture, by identifying Cartier's "cosmic Galois group" as a universal group of symmetries that organizes the structure of the divergences in perturbative quantum field theory. We show that the group can also be realized as a motivic Galois group, and has therefore precisely the type of arithmetic nature expected by Cartier.

This group acts on the constants of physical theories. In fact, it maps to the group of diffeographisms of any given renormalizable theory by a representation determined by the beta function of the theory. Therefore, it also maps to the group of formal diffeomorphisms of the coupling constants by the results of §6.5 above.

We will also see that the renormalization group sits naturally as a 1-parameter subgroup of this universal group and this provides an interpretation of the renormalization group as Galois symmetries.

The main steps, which we are going to discuss in detail in the rest of this section, are summarized as follows.

- The data of counterterms as iterated integrals (Gross–'t Hooft relations).
- A geometric formulation: flat equisingular connections.
- The local Riemann–Hilbert correspondence: differential Galois group through the Tannakian formalism.
- The universal singular frame and universal symmetries.

7.1. Counterterms and time-ordered exponentials.

The first result we describe here (Theorem 1.58 below) is a refinement of Theorem 2 of [**83**], which proved the analog in the context of the CK theory of a well known result of D. Gross ([**151**] §4.5) and 't Hooft (unpublished). The latter states that the counterterms in perturbative renormalization only depend on the beta function of the theory (Gross–'t Hooft relations).

We first need to discuss the mathematical formulation of the *time-ordered exponential.*

We assume that $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$ is a positively graded connected commutative Hopf algebra over $\mathbb{C}$, with G its affine group scheme. We have $\mathfrak{g} = \operatorname{Lie} G$ as in Definition 1.21. When the $\mathcal{H}_n$ are finite-dimensional vector spaces, the Lie algebra $\mathfrak{g}$ is related to the dual Hopf algebra $\mathcal{H}^\vee$ as in Theorem 1.22.

Let H be a Lie group and $t \mapsto \alpha(t)$ a smooth map from the interval $[a, b]$ to the Lie algebra of H.

Dyson's time-ordered exponential $\mathrm{T}e^{\int_a^b \alpha(t)\,dt}$ is a notation for the parallel transport in the trivial H-principal bundle $[a, b] \times H$ with the left action of H, endowed with the connection associated to the 1-form $\alpha(t)\,dt$ with values in the Lie algebra of H. In other words $h(u) = \mathrm{T}e^{\int_a^u \alpha(t)\,dt} \in H$ is the solution of the differential equation

$$dh(u) = h(u)\,\alpha(u)du\,, \quad h(a) = 1\,. \tag{1.261}$$

A mathematical definition can be given as an iterated integral. This type of formalism was developed in the topological context in [**55**], [**56**] (Chen's iterated integral) and in the operator algebra context in [**3**] (Araki's expansional). We give here a formulation adapted to the context of affine group schemes.

DEFINITION 1.50. *Given a $\mathfrak{g}(\mathbb{C})$-valued smooth function $\alpha(t)$, with $t \in [a, b] \subset \mathbb{R}$, the time-ordered exponential (also called the expansional) is defined as*

$$\mathrm{T}e^{\int_a^b \alpha(t)\,dt} := 1 + \sum_1^\infty \int_{a\leq s_1\leq\cdots\leq s_n\leq b} \alpha(s_1)\cdots\,\alpha(s_n)\;ds_1\cdots ds_n\,, \tag{1.262}$$

with the product taken in $\mathcal{H}^\vee$, and with $1 \in \mathcal{H}^\vee$ the unit corresponding to the counit ε of $\mathcal{H}$.

One has the following result, which in particular shows that the expansional only depends on the 1-form $\alpha(t)dt$.

PROPOSITION 1.51. *The time-ordered exponential* (1.262) *satisfies the following properties:*

(1) *When paired with any $X \in \mathcal{H}$ the sum* (1.262) *is finite.*
(2) *The iterated integral* (1.262) *defines an element of $G(\mathbb{C})$.*
(3) *The expansional* (1.262) *is the value $g(b)$ of the unique solution $g(t) \in G(\mathbb{C})$ with initial condition $g(a) = 1$ of the differential equation*

$$dg(t) = g(t)\,\alpha(t)\,dt\,. \tag{1.263}$$

(4) *The iterated integral is multiplicative over the sum of paths:*

$$\mathrm{T}e^{\int_a^c \alpha(t)\,dt} = \mathrm{T}e^{\int_a^b \alpha(t)\,dt}\,\mathrm{T}e^{\int_b^c \alpha(t)\,dt}. \tag{1.264}$$

(5) *Let $\rho : G(\mathbb{C}) \to H$ be a homomorphism to a Lie group H; then $\rho(\mathrm{T}e^{\int_a^b \alpha(t)\,dt})$ is the parallel transport in the principal bundle $[a, b] \times H$ endowed with the connection associated to the 1-form $\rho(\alpha(t))\,dt$.*
(6) *The inverse in $G(\mathbb{C})$ of the element* (1.262) *is of the form*

$$(\mathrm{T}e^{\int_a^b \alpha(t)\,dt})^{-1} = \mathrm{T}'e^{-\int_a^b \alpha(t)\,dt}, \tag{1.265}$$

with

$$\mathrm{T}'e^{\int_a^b \alpha(t)\,dt} := 1 + \sum_1^\infty \int_{a\leq s_1\leq\cdots\leq s_n\leq b} \alpha(s_n)\cdots\,\alpha(s_1)\;ds_1\cdots ds_n\,. \tag{1.266}$$

PROOF. (1) The elements $\alpha(t) \in \mathfrak{g}(\mathbb{C})$, viewed as linear forms on $\mathcal{H}$, vanish on any element of degree 0. Thus, for $X \in \mathcal{H}$ of degree n, one has

$$\langle \alpha(s_1) \cdots \alpha(s_m),\, X \rangle \,=\, 0 \quad \forall m > n\,,$$

so that

$$\langle \mathrm{T} e^{\int_a^b \alpha(t)\,dt}, X \rangle \tag{1.267}$$

is a finite sum.

(2) First notice that (1.267) satisfies

$$\partial_t \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X \rangle = \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}\, \alpha(t), X \rangle. \tag{1.268}$$

In fact, differentiating each term in the finite sum (1.267) in the variable t amounts to fixing the last integration variable $s_n = t$ in

$$\int_{a \leq s_1 \leq \cdots \leq s_n \leq t} \langle \alpha(s_1) \cdots \alpha(s_n),\, X \rangle \; ds_1 \cdots ds_n.$$

In order to show that (1.262) defines an element in $G(\mathbb{C})$, we need to check that for all $X, Y \in \mathcal{H}$,

$$\langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, XY \rangle = \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X \rangle \; \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, Y \rangle. \tag{1.269}$$

We show this for homogeneous X and Y, by induction on the sum of their degrees. Using (1.268), we write

$$\begin{aligned} \partial_t \, \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X\,Y \rangle = & \;\; \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}\, \alpha(t), X\,Y \rangle \\ = & \;\; \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds} \otimes \alpha(t), \Delta(X)\,\Delta(Y) \rangle\,. \end{aligned}$$

We then write the coproduct in the form

$$\Delta(X) = X_{(1)} \otimes X_{(2)} = X \otimes 1 + 1 \otimes X + \sum X' \otimes X''$$

where only terms of lower degree appear in the last sum. Using the derivation property (1.173) of $\alpha(t) \in \mathfrak{g}(\mathbb{C})$, one gets that it pairs trivially with all the products $X_{(2)}\, Y_{(2)}$ except when $Y_{(2)} = 1$ or $X_{(2)} = 1$. This gives

$$\begin{aligned} \partial_t \, \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X\,Y \rangle = & \;\; \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X_{(1)}\,Y \rangle \; \langle \alpha(t), X_{(2)} \rangle \\ & + \;\; \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X\,Y_{(1)} \rangle \; \langle \alpha(t), Y_{(2)} \rangle\,. \end{aligned}$$

We can then apply the induction hypothesis and obtain

$$\begin{aligned} \partial_t \, \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X\,Y \rangle \;\; & = \partial_t \left(\langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X \rangle \right) \; \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, Y \rangle \\ & + \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X \rangle \; \partial_t \left(\langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, Y \rangle \right). \end{aligned}$$

If we denote

$$F(t) := \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X\,Y \rangle - \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, X \rangle \; \langle \mathrm{T} e^{\int_a^t \alpha(s)\,ds}, Y \rangle$$

we obtain $\partial_t F(t) = 0$, for all t. Thus, since for $t = a$ we clearly have $F(a) = 0$, we obtain (1.269).

(3) This then follows immediately from (1.268).

(4) The property (1.264) follows from (3), since both sides satisfy equation (1.263) and agree for $c = b$, hence they are equal.

(5) This follows from (3) and (1.261).
(6) Notice that one can apply the notion of expansional to the opposite algebra. This amounts to reversing the order of the terms. Thus, (1.262) gets replaced by (1.266). As in (3) above, one gets that, for $t > 0$, the expression $g(t) = \mathrm{T}' e^{\int_0^t \alpha(u)\, du}$ is the unique solution to the ODE

$$dg(t) = \alpha(t)\, g(t)\, dt\,. \tag{1.270}$$

This shows that the inverse of a time-ordered exponential with T is given by the time-ordered exponential with T', which gives (1.265). ☐

Consider now an open domain $\Omega \subset \mathbb{R}^2$ and, for $(s,t) \in \Omega$, let

$$\varpi = \alpha(s,t)ds + \eta(s,t)dt \tag{1.271}$$

be a *flat* $\mathfrak{g}(\mathbb{C})$-valued connection. The flatness condition means that we have

$$\partial_s\, \eta - \partial_t\, \alpha + [\alpha, \eta] = 0. \tag{1.272}$$

The time-ordered exponential satisfies the following property

PROPOSITION 1.52. *Let ϖ be a flat $\mathfrak{g}(\mathbb{C})$-valued connection on $\Omega \subset \mathbb{R}^2$ as above, and let $\gamma : [0,1] \to \Omega$ be a path in Ω. Then the time-ordered exponential $\mathrm{T}e^{\int_0^1 \gamma^* \varpi}$ only depends on the homotopy class $[\gamma]$ of paths with endpoints $a = \gamma(0)$ and $b = \gamma(1)$.*

This follows from Proposition 1.51, using the fact that $G(\mathbb{C})$ is a projective limit of Lie groups.

Recall that a differential field (K, δ) is a field K endowed with a derivation satisfying $\delta(f+h) = \delta(f) + \delta(h)$ and $\delta(fh) = f\delta(h) + \delta(f)h$, for all $f, h \in K$. The field of constants is the subfield $\{f \in K | \delta(f) = 0\}$.

DEFINITION 1.53. *Let (K, δ) be a differential field with differentiation $f \mapsto f' := \delta(f)$ and with field of constants $\mathbb{C}$. The logarithmic derivative on the group $G(K) = \mathrm{Hom}_{\mathcal{A}_\mathbb{C}}(\mathcal{H}, K)$ is given by*

$$D(g) := g^{-1}\, g' \in \mathfrak{g}(K), \quad \forall g \in G(K), \tag{1.273}$$

where $g' = \delta(g)$ is the linear map $g' : \mathcal{H} \to K$,

$$g'(X) = \delta(g(X)), \quad \forall X \in \mathcal{H}.$$

It is not hard to check that (1.273) indeed defines an element in $\mathfrak{g}(K)$. In fact, $D(g) : \mathcal{H} \to K$ is the linear map

$$\langle D(g), X \rangle = g^{-1} \star g'\,(X) = \langle g^{-1} \otimes g', \Delta X \rangle$$

and this satisfies

$$\langle D(g), X\, Y \rangle = \langle D(g), X \rangle\, \varepsilon(Y) + \varepsilon(X)\, \langle D(g), Y \rangle, \quad \forall X, Y \in \mathcal{H}.$$

In the following, we assume that $K = \mathbb{C}(\{z\})$, the differential field of convergent Laurent series, as in (1.217), with $\delta(f) = \frac{d}{dz} f$. Later, we will also consider the case where $K = \mathbb{C}((z)) = \mathbb{C}[[z]][z^{-1}]$ is the field of formal Laurent series.

LEMMA 1.54. *Let $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$ be a positively graded connected commutative Hopf algebra over $\mathbb{C}$, with G its affine group scheme. Let $\mathcal{H}(i) \subset \mathcal{H}$ be an increasing sequence of finitely generated commutative Hopf subalgebras, with $\mathcal{H} = \cup\mathcal{H}(i)$. The affine group scheme G is the projective limit (1.171) of the linear algebraic groups G_i dual to the finitely generated commutative Hopf subalgebras $\mathcal{H}(i) \subset \mathcal{H}$. Suppose given an element $\varpi \in \mathfrak{g}(K)$, for $K = \mathbb{C}(\{z\})$.*

(1) *There is a well defined monodromy representation*

$$M_i(\varpi) : \mathbb{Z} \to G_i(\mathbb{C}).$$

(2) *The condition of trivial monodromy $M(\varpi) = 1$ is well defined in $G(\mathbb{C})$.*

PROOF. (1) For $\varpi \in \mathfrak{g}(K)$, and a fixed G_i, consider the corresponding element $\varpi_i \in \mathfrak{g}_i(K)$ obtained by restriction to $\mathcal{H}(i) \subset \mathcal{H}$. Since $\mathcal{H}(i)$ is finitely generated, ϖ_i depends on only finitely many elements in K, hence there is a radius $\rho_i > 0$ such that all these finitely many Laurent series elements converge in a punctured disk Δ_i^* of this radius around $z = 0$. Thus, we can choose a base point $z_i \in \Delta_i^*$. Composing with the evaluation $f \to f(z)$ for $z \in \Delta_i^*$ we get a flat $\mathfrak{g}_i(\mathbb{C})$-valued connection $\omega_i = \varpi_i(z)dz$ on Δ_i^*. For any smooth path $\gamma : [0,1] \to \Delta_i^*$ the expression

$$M_i(\varpi)(\gamma) := \mathrm{T}e^{\int_0^1 \gamma^*\omega_i} \in G_i(\mathbb{C}) \tag{1.274}$$

is well defined. By Proposition 1.52 and the flatness of the connection ω_i, viewed as a connection in two real variables, (1.274) depends only on the homotopy class of the loop γ in Δ_i^* with $\gamma(0) = \gamma(1) = z_i$. Thus, it defines a monodromy representation

$$M_i(\varpi) : \pi_1(\Delta_i^*, z_i) \to G_i(\mathbb{C}). \tag{1.275}$$

(2) By construction, the conjugacy class of M_i does not depend on the choice of the base point. When passing to the projective limit, one has to take care of the change of base points, but the condition of *trivial monodromy*

$$M(\varpi) = 1 \Leftrightarrow M_i(\varpi) = 1\,, \quad \forall i \tag{1.276}$$

is well defined at the level of the projective limit G of the groups G_i.

□

The monodromy encodes the obstruction to the existence of solutions to the differential equation $D(g) = \varpi$, for $g \in \mathfrak{g}(K)$.

EXAMPLE 1.55. *Let $\mathbb{G}_a$ be the additive group. Let as above $K = \mathbb{C}(\{z\})$ with $\delta(f) = \frac{d}{dz}f$. One has $\mathbb{G}_a(K) = K$, $D(f) = \delta(f) = f'$, and the residue of $\varpi \in K$ is a non-trivial obstruction to the existence of solutions of the equation $D(f) = \varpi$.*

The following result shows that the monodromy is in fact the only obstruction.

PROPOSITION 1.56. *Let $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$ be a positively graded connected commutative Hopf algebra over $\mathbb{C}$, with G its affine group scheme. Suppose given $\varpi \in \mathfrak{g}(K)$ with trivial monodromy, $M(\varpi) = 1$. Then there exists a solution $g \in G(K)$ of the equation*

$$D(g) = \varpi\,. \tag{1.277}$$

PROOF. We view, as above, G as the projective limit of the G_i and use the same notations as above. We treat the case of a fixed G_i first. We let

$$g_i(z) = \mathrm{T}e^{\int_{z_i}^{z} \omega_i}\,, \tag{1.278}$$

where by the right-hand side we mean the time-ordered exponential $\mathrm{T}e^{\int_0^1 \gamma^*\omega_i}$ for any smooth path in Δ_i^* from the base point z_i to z. The condition of trivial monodromy ensures that the result does not depend on the choice of the path γ. Let $\mathcal{H}(i)$ be as in Lemma 1.54. One needs to show that, for any $X \in \mathcal{H}(i) \subset \mathcal{H}$, the evaluation

$$h_i(z) = \langle g_i(z),\, X\rangle$$

is a convergent Laurent series in Δ_i^*. By construction the same property holds for $\varpi_i(z)$. Moreover, by Proposition 1.51 (1) we know that, when pairing with $X \in \mathcal{H}(i)$ only finitely many terms of the infinite sum (1.262) defining $g_i(z)$ contribute. Thus, it follows that $z^{N_i}\, h_i(z)$ is bounded in Δ_i^*, for sufficiently large N_i. Moreover, by (1.263) of Proposition 1.51, with $2\bar\partial = \partial_t + i\partial_s$ and $z = t + is$, we obtain $\overline{\partial} h_i = 0$, which gives $h_i \in K$. Thus $g_i \in G_i(K)$.

Finally (1.263) of Proposition 1.51 shows that $g_i \in G_i(K)$ is a solution of

$$D(g_i) = \varpi_i\,. \tag{1.279}$$

Given two solutions g and h of (1.279), one gets that $g\,h^{-1}$ satisfies $D(gh^{-1}) = 0$. Thus since the field of constants of K is $\mathbb{C}$, it follows that there exists an element $a \in G_i(\mathbb{C})$ such that $g = ah$. Now since the $\mathcal{H}(i) \subset \mathcal{H}$ are Hopf subalgebras of $\mathcal{H}$, the canonical projection

$$p_i : G_{i+1}(\mathbb{C}) \to G_i(\mathbb{C})$$

is surjective. Thus, in order to pass to the projective limit, one constructs by induction a projective system of solutions $g_i \in G_i(K)$, where one modifies the solution in $G_{i+1}(K)$ by left multiplication by an element of $G_{i+1}(\mathbb{C})$ so that it projects onto g_i. $\square$

Example 1.57. *It is crucial in Proposition* 1.56 *to assume that* $\mathcal{H}$ *is positively graded connected. Let* $\mathbb{G}_m$ *be the multiplicative group. Let as above* $K = \mathbb{C}(\{z\})$ *with* $\delta(f) = \frac{d}{dz} f$. *One has* $\mathbb{G}_m(K) = K^*$, $D(f) = f^{-1}\,\delta(f)$, *and with* $\varpi = \frac{1}{z^2} \in K$ *the equation* $D(f) = \varpi$ *has trivial monodromy with formal solution given by*

$$f(z) = e^{-1/z}$$

but no solution in $\mathbb{G}_m(K) = K^*$.

With these general results about time-ordered exponentials in place, we come to the Gross–'t Hooft relations. In our context, the fact that the counterterms in perturbative renormalization depend only on the beta function is a consequence of the following explicit form of the negative piece of the Birkhoff factorization as a time-ordered exponential.

Theorem 1.58. *Suppose given* $\gamma_\mu(z) \in L(G(\mathbb{C}), \mu)$. *Then the negative piece* $\gamma_-(z)$ *of the Birkhoff factorization has the explicit form*

$$\gamma_-(z) = \mathrm{T}e^{-\frac{1}{z}\int_0^\infty \theta_{-t}(\beta)\,dt}. \tag{1.280}$$

Proof. We write $\gamma_-(z)^{-1} = 1 + \sum_{n=1}^\infty \frac{d_n}{z^n}$, as in (1.252) of Lemma 1.48. To obtain the result we first show that the coefficients d_n are given by the formula (cf. [**83**])

$$d_n = \int_{s_1 \geq s_2 \geq \cdots \geq s_n \geq 0} \theta_{-s_1}(\beta)\, \theta_{-s_2}(\beta) \cdots \theta_{-s_n}(\beta)\; ds_1 \cdots ds_n\,. \tag{1.281}$$

To check this, notice first that, for $n = 1$, the expression (1.281) reduces to

$$d_1 = \int_0^\infty \theta_{-s}(\beta)\, ds\,. \tag{1.282}$$

This follows from the equality

$$Y^{-1}(X) = \int_0^\infty \theta_{-s}(X)\, ds, \tag{1.283}$$

for all $X \in \mathcal{H}$ with $\varepsilon(X) = 0$, where ε is the counit of $\mathcal{H}$. In fact, given α and α' in $\mathcal{H}^\vee$ satisfying $\alpha' = Y(\alpha)$ and $\langle \alpha, 1\rangle = \langle \alpha', 1\rangle = 0$, the identity (1.283) implies

$$\alpha = \int_0^\infty \theta_{-s}(\alpha')\, ds\,.$$

This gives (1.282) when applied to $\beta = Y(d_1)$ of Corollary 1.49. The general formula (1.281) then follows by induction, applying the relations between the coefficients d_n proved in Lemma 1.48 and the fact that the 1-parameter family θ_t generated by the grading of $\mathcal{H}$ also acts as automorphisms of $\mathcal{H}^\vee$.

To get (1.280) from (1.281), we use (6) of Proposition 1.51. Using (1.281) we get

$$\gamma_-(z)^{-1} = \mathrm{T}' e^{\frac{1}{z}\int_0^\infty \theta_{-s}(\beta)\, ds}$$

and the required equality (1.280) follows from (1.265). ☐

Remark 1.59. *One does not have to worry about the convergence issue in dealing with the iterated integrals* (1.281) *in* (1.280).

In fact, when evaluating (1.281) on elements $X \in \mathcal{H}$, one obtains for $\langle d_n, X\rangle$ the expression

$$\int_{s_1 \geq \cdots \geq s_n \geq 0} \langle \beta \otimes \cdots \otimes \beta, \theta_{-s_1}(X_{(1)}) \otimes \theta_{-s_2}(X_{(2)}) \otimes \cdots \otimes \theta_{-s_n}(X_{(n)})\rangle ds_1 \cdots ds_n,$$

where we used the notation

$$\Delta^{(n-1)}(X) = \sum X_{(1)} \otimes X_{(2)} \otimes \cdots \otimes X_{(n)}\,.$$

The convergence of the iterated integral is exponential. In fact, we have the estimate $\langle \beta, \theta_{-s}(X_{(i)})\rangle = O\left(e^{-s}\right)$ for $s \to +\infty$. If X is homogeneous of degree $\deg(x)$ and if $n > \deg(x)$, then at least one of the $X_{(i)}$ has degree 0, and $\langle \beta, \theta_{-s}(X_{(i)})\rangle = 0$ so that $\langle d_n, X\rangle$ vanishes. This implies that, when pairing $\gamma_-(z)^{-1}$ with $X \in \mathcal{H}$, only finitely many terms contribute to

$$\langle \gamma_-(z)^{-1}, X\rangle = \varepsilon(X) + \sum_{n=1}^\infty \frac{1}{z^n}\, \langle d_n, X\rangle\,.$$

This remark will be useful when we compare the non-formal case, where $K = \mathbb{C}(\{z\})$, with the formal case, where $K = \mathbb{C}((z))$ (cf. §7.6 below).

7.2. Flat equisingular connections.

The main advantage of introducing the formalism of time-ordered exponentials as iterated integrals, as in Definition 1.50, is their property of being characterized as solutions of a differential equation (1.263), as in Proposition 1.51. In fact, this makes it possible to translate data described by time-ordered exponentials into a class of differential equations.

Following this idea, we obtain in this section a geometric characterization of the divergences in renormalizable quantum field theory, in terms of a suitable class of differential systems.

As a first step, we give a more precise characterization of the class $L(G(\mathbb{C}),\mu)$ of loops $\gamma_\mu(z)$ satisfying the two conditions (1.240) and (1.241), as in Definition 1.45. For instance, given any loop $\gamma_{\rm reg}(z)$ which is regular at $z=0$, one can easily obtain an element $\gamma_\mu \in L(G(\mathbb{C}),\mu)$ by setting $\gamma_\mu(z)=\theta_{z\log\mu}(\gamma_{\rm reg}(z))$. The following result shows that the only remaining piece of data that is necessary, in order to specify completely the loops that belong to the class $L(G(\mathbb{C}),\mu)$, is the choice of an element β in the Lie algebra of the affine group scheme G.

Theorem 1.60. *Suppose given a loop $\gamma_\mu(z)$ in the class $L(G(\mathbb{C}),\mu)$. Then the following properties hold.*

(1) *The loop $\gamma_\mu(z)$ has the form*

$$\gamma_\mu(z) = \mathrm{T}e^{-\frac{1}{z}\int_\infty^{-z\log\mu}\theta_{-t}(\beta)\,dt}\,\theta_{z\log\mu}(\gamma_{\rm reg}(z)), \tag{1.284}$$

for a unique $\beta\in\mathfrak{g}(\mathbb{C})$, with $\gamma_{\rm reg}(z)$ a loop regular at $z=0$.

(2) *The Birkhoff factorization of $\gamma_\mu(z)$ has the form*

$$\begin{aligned}\gamma_{\mu^+}(z) &= \mathrm{T}e^{-\frac{1}{z}\int_0^{-z\log\mu}\theta_{-t}(\beta)\,dt}\,\theta_{z\log\mu}(\gamma_{\rm reg}(z))\,,\\ \gamma_-(z) &= \mathrm{T}e^{-\frac{1}{z}\int_0^{\infty}\theta_{-t}(\beta)\,dt}\,.\end{aligned} \tag{1.285}$$

(3) *Conversely, given an element $\beta\in\mathfrak{g}(\mathbb{C})$ and a regular loop $\gamma_{\rm reg}(z)$, the expression* (1.284) *gives an element $\gamma_\mu\in L(G(\mathbb{C}),\mu)$.*

Proof. (1) Let $\gamma_\mu(z)$ be a loop in $L(G(\mathbb{C}),\mu)$, with Birkhoff factorization $\gamma_\mu(z) = \gamma_-(z)^{-1}\gamma_{\mu^+}(z)$. Consider the loop

$$\alpha_\mu(z) := \theta_{z\log\mu}(\gamma_-(z)^{-1}). \tag{1.286}$$

This satisfies the scaling property (1.240) by construction,

$$\alpha_{e^s\mu}(z) = \theta_{sz}(\alpha_\mu(z)). \tag{1.287}$$

Consider the ratio $\alpha_\mu(z)^{-1}\gamma_\mu(z)$. This still satisfies the scaling property (1.240). Moreover, the loop $\alpha_\mu(z)^{-1}\gamma_\mu(z)$ is regular at $z=0$. In fact, this holds for $\mu=1$ and hence all values of μ by the scaling property (1.240). Thus, it can be written in the form

$$\alpha_\mu(z)^{-1}\gamma_\mu(z) = \theta_{z\log\mu}(\gamma_{\rm reg}(z)), \tag{1.288}$$

with $\gamma_{\rm reg}(z)$ a regular loop.

Notice then that, using (1.286) and (1.280) of Theorem 1.58, we obtain

$$\alpha_\mu(z)^{-1} = \mathrm{T}e^{-\frac{1}{z}\int_{-z\log\mu}^{\infty}\theta_{-t}(\beta)\,dt}. \tag{1.289}$$

Thus, using (1.264), we obtain (1.284).

(2) Given $\gamma_\mu \in L(G(\mathbb{C}),\mu)$, we write it in the form (1.284) as above. Thus, we have

$$\gamma_\mu(z)^{-1} = \theta_{z\log\mu}(\gamma_{\mathrm{reg}}(z))^{-1}\mathrm{T}e^{-\frac{1}{z}\int_{-z\log\mu}^{\infty}\theta_{-t}(\beta)\,dt}.$$

We write the second term, using (1.264), as the product

$$\mathrm{T}e^{-\frac{1}{z}\int_{-z\log\mu}^{0}\theta_{-t}(\beta)\,dt}\,\mathrm{T}e^{-\frac{1}{z}\int_{0}^{\infty}\theta_{-t}(\beta)\,dt}.$$

Thus, we have obtained the expression

$$\gamma_\mu(z)^{-1} = \theta_{z\log\mu}(\gamma_{\mathrm{reg}}(z))^{-1}\ \mathrm{T}e^{-\frac{1}{z}\int_{-z\log\mu}^{0}\theta_{-t}(\beta)\,dt}\gamma_-(z),$$

where $\gamma_-(z)$ is a regular function of $1/z$ with $\gamma_-(\infty) = 1$. We then need to check the regularity at $z = 0$ of the remaining part

$$\mathrm{T}e^{-\frac{1}{z}\int_{0}^{-z\log\mu}\theta_{-t}(\beta)\,dt}\theta_{z\log\mu}(\gamma_{\mathrm{reg}}(z)).$$

It is enough to show that we have

$$\lim_{z\to 0}\mathrm{T}e^{-\frac{1}{z}\int_{0}^{-sz}\theta_{-t}(\beta)\,dt} = e^{s\beta}. \tag{1.290}$$

To see this we take the straight path from 0 to $-sz$ in the form $\pi(u) = -szu$ for $u \in [0,1]$ and write

$$\mathrm{T}e^{-\frac{1}{z}\int_{0}^{-sz}\theta_{-t}(\beta)\,dt} = \mathrm{T}e^{\int_0^1\alpha(u)\,du}$$

where the form $\alpha(u)\,du$ is the pullback by π of the form $-\frac{1}{z}\,\theta_{-t}(\beta)\,dt$. So far both s and z are fixed. One has

$$\alpha(u)\,du = -\frac{1}{z}\,\theta_{szu}(\beta)\,(-sz)\,du = s\,\theta_{szu}(\beta)\,du$$

This gives

$$\mathrm{T}e^{-\frac{1}{z}\int_{0}^{-sz}\theta_{-t}(\beta)\,dt} = \mathrm{T}e^{s\int_0^1\theta_{szu}(\beta)\,du}$$

which converges when $z \to 0$ to $e^{s\beta}$. We then obtain the Birkhoff factorization in the form (1.285).

(3) Given a choice of an element $\beta \in \mathfrak{g}(\mathbb{C})$ and a regular loop $\gamma_{\mathrm{reg}}(z)$, the expression (1.284) satisfies the scaling property (1.240) by construction. The formula (1.285) for the Birkhoff factorization also shows clearly that the negative piece is independent of μ so that (1.241) is also satisfied, hence $\gamma_\mu \in L(G(\mathbb{C}),\mu)$. ☐

In the rest of this section, we use the classification result of Theorem 1.60, together with the properties of the time-ordered exponential, to give a geometric reformulation of the class $L(G(\mathbb{C}),\mu)$ in terms of a class of differential systems associated to a family of singular flat connections.

We introduce some notation. Let (K,δ) be a differential field with field of constants $\mathbb{C}$. We will assume $K = \mathbb{C}(\{z\})$ with $\delta(f) = f'$. We have $\mathcal{O} = \mathbb{C}\{z\}$, the subring of convergent power series as in (1.219).

We denote by $\Omega^1(\mathfrak{g})$ the space of 1-forms on K with values in $\mathfrak{g}$. These are all of the form $A(z)\,dz$ for some $A \in \mathfrak{g}(K)$. The differential

$$d : K \to \Omega^1$$

is given by $df = \delta(f)\,dz = \frac{df}{dz}\,dz$.

The logarithmic derivative D on $G(K)$, as in (1.273) of Definition 1.53, induces an operator (which, with a slight abuse of notation, we still call D)

$$D : G(K) \to \Omega^1(\mathfrak{g}), \qquad Df = f^{-1}\, df. \tag{1.291}$$

This satisfies the property

$$D(fh) = Dh + h^{-1}\, Df\, h. \tag{1.292}$$

One can then consider differential equations of the form

$$Df = \varpi \tag{1.293}$$

for $\varpi \in \Omega^1(\mathfrak{g})$.

We can think of the element $\varpi \in \Omega^1(\mathfrak{g})$ as a connection on the trivial principal G-bundle over an infinitesimal punctured disk Δ^* around $z = 0$. In fact, a connection on the trivial principal G-bundle $\Delta^* \times G$ is specified by the restriction of the connection form to $\Delta^* \times 1$, i.e. by a $\mathfrak{g}$-valued 1-form ϖ on Δ^*, which is the same as an element in $\Omega^1(\mathfrak{g})$. We use the principal bundle point of view in §7.3 below. For the moment, we simply define a connection as a $\mathfrak{g}$-valued 1-form ϖ.

We then have the following natural equivalence relation for connections $\varpi \in \Omega^1(\mathfrak{g})$.

Definition 1.61. *Two connections ϖ and ϖ' are equivalent iff they are gauge conjugate by an element regular at $z = 0$, i.e.*

$$\varpi' = Dh + h^{-1} \varpi\, h, \qquad \textit{with} \quad h \in G(\mathcal{O}). \tag{1.294}$$

As in Lemma 1.54, the condition of trivial monodromy $M(\varpi) = 1$ is well defined and it ensures the existence of solutions to (1.293). Suppose given a solution $f \in G(K)$ to (1.293). We know, by the same argument as used in Theorem 1.39, that loops $f \in G(K)$ have a unique Birkhoff factorization

$$f = f_-^{-1} f_+, \tag{1.295}$$

with $f_+ \in G(\mathcal{O})$ and $f_- \in G(\tilde{\mathcal{Q}})$ satisfying $\varepsilon_- \circ f_- = \varepsilon$ as in Definition 1.38.

The following result then explains in more detail the meaning of the equivalence relation of Definition 1.61.

Proposition 1.62. *Suppose given two connections ϖ and ϖ' in $\Omega^1(\mathfrak{g})$ with trivial monodromy. Then ϖ and ϖ' are equivalent, in the sense of* (1.294)*, iff there exist solutions to the corresponding equations* (1.293) *that have the same negative pieces of the Birkhoff factorization,*

$$\varpi \sim \varpi' \Longleftrightarrow f_-^{\varpi} = f_-^{\varpi'}, \tag{1.296}$$

for some f^{ϖ} and $f^{\varpi'}$ in $G(K)$ satisfying $D(f^{\varpi}) = \varpi$ and $D(f^{\varpi'}) = \varpi'$.

Proof. By Lemma 1.54, we know that there exist solutions, since the monodromy is trivial. We first show that ϖ is equivalent to the 1-form $D((f_-^{\varpi})^{-1})$. We have $f^{\varpi} = (f_-^{\varpi})^{-1} f_+^{\varpi}$, hence the product rule (1.292) gives the required equivalence, since $f_+^{\varpi} \in G(\mathcal{O})$. This shows that, if $f_-^{\varpi} = f_-^{\varpi'}$ for some choices of solutions, then ϖ and ϖ' are equivalent.

Conversely, assume that ϖ and ϖ' are equivalent, i.e. $\varpi' = Dh + h^{-1}\varpi\, h$. We take a solution f^{ϖ} and let $f^{\varpi'}$ be given by

$$f^{\varpi'} = f^{\varpi}\, h\,,$$

for the above $h \in G(\mathcal{O})$. It follows from (1.292) that $D(f^{\varpi'}) = \varpi'$. The uniqueness of the Birkhoff factorization then gives $f_-^{\varpi} = f_-^{\varpi'}$. ☐

Definition 1.63. *Two elements f_j of $G(K)$ have the same singularity iff*

$$f_1^{-1}\, f_2 \in G(\mathcal{O}) \tag{1.297}$$

Equivalently this means that the negative parts of their Birkhoff decompositions are the same.

From the point of view of renormalization, introducing the equivalence relation of Definitions 1.61 and 1.63 means that one is interested in retaining only the information on the behavior of the divergences, which is encoded in the counterterms given by the negative piece $\gamma_-(z)$ of the Birkhoff factorization of Proposition 1.40.

In order to make this geometric setting suitable to treat the data of perturbative renormalization, i.e. the class of loops $L(G(\mathbb{C}), \mu)$, we need to account for the mass parameter μ in the geometry.

This can be done by considering a principal $\mathbb{G}_m(\mathbb{C}) = \mathbb{C}^*$-bundle B,

$$\mathbb{G}_m \to B \xrightarrow{\pi} \Delta\,, \tag{1.298}$$

over the infinitesimal disk Δ. We let $P = B \times G$ be the trivial principal G-bundle over the base space B.

We denote by V the fiber over $z = 0$,

$$V = \pi^{-1}(\{0\}) \subset B,$$

and by B^0 its complement

$$B^0 = B \setminus V \subset B.$$

We also fix a base point $y_0 \in V$. We let $P^0 = B^0 \times G$ denote the restriction to B^0 of the bundle P.

Remark 1.64. The physical meaning of these data is the following. In the case of a renormalizable theory $\mathcal{T}$, the group scheme is $G = \mathrm{Difg}(\mathcal{T})$. The disk Δ corresponds to the complexified dimension $z \in \Delta$ of DimReg. The principal bundle B over Δ encodes all the possible choices of normalization for the "integral in dimension $D-z$" of (1.135), hence it accounts for the presence of the mass parameter μ. The choice of a base point $y_0 \in V$ corresponds to fixing the value of the Planck constant $\hbar$. One needs to be careful and refrain from considering the fiber $\pi^{-1}(\{z\}) \subset B$ as the set of possible values of μ. It is rather the set of possible values of $\mu^z\,\hbar$. In fact in all the above discussion we restricted to $\mu \in \mathbb{R}_+^*$ and the fact that $\log\mu$ makes sense did play a role. The role of the choice of a unit of mass μ is the same as the choice of a section $\sigma : \Delta \to B$ (up to order one).

We have an action of $\mathbb{G}_m$ on B, which we write as

$$b \mapsto u(b), \quad \forall u \in \mathbb{G}_m(\mathbb{C}) = \mathbb{C}^*\,.$$

Since the affine group scheme G is dual to a graded Hopf algebra $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$, we also have an action of $\mathbb{G}_m$ on G, as in Lemma 1.25, induced by the action u^Y of

(1.176) on the Hopf algebra, where Y is the grading operator on $\mathcal{H}$. We denote the action of $\mathbb{G}_m$ on G also by u^Y, as the meaning is clear. Thus, we have a $\mathbb{G}_m$-action on $P = B \times G$ given by

$$u(b, g) = (u(b), u^Y(g)), \quad \forall u \in \mathbb{G}_m . \tag{1.299}$$

We explain in §7.3 below how this action, which as such is not given by automorphisms of a G-bundle, in fact comes from a $\mathbb{G}_m$-equivariant principal bundle over B.

We now introduce the class of connections that will be our main object of interest, as they encode geometrically the properties of the class of loops $L(G(\mathbb{C}), \mu)$.

Definition 1.65. *We say that a flat connection ϖ on P^0 is* equisingular *iff ϖ is $\mathbb{G}_m$-invariant and for any solution γ, $D\gamma = \varpi$, the restrictions of γ to sections $\sigma : \Delta \to B$ with $\sigma(0) = y_0$ have the same singularity.*

We can fix a choice of a (non-canonical) regular section $\sigma : \Delta \to B$ of the fibration (1.298), with $\sigma(0) = y_0$. Then the first condition of Definition 1.65 can be written in the form

$$\varpi(z, u(v)) = u^Y(\varpi(z, v)), \quad \forall u \in \mathbb{G}_m, \tag{1.300}$$

with $v = (\sigma(z), g)$, for $z \in \Delta$ and $g \in G$.

The second condition states that, if σ_1 and σ_2 are two sections of B as above, with $\sigma_1(0) = y_0 = \sigma_2(0)$, then

$$\sigma_1^*(\gamma) \sim \sigma_2^*(\gamma), \tag{1.301}$$

that is, the pullbacks of γ by these sections are equivalent on $\Delta^* \times G$, through the equivalence (1.297) of Definition 1.63. This means that $\sigma_1^*(\gamma)$ and $\sigma_2^*(\gamma)$ have the same singularity at the origin $z = 0$, hence the use of the term "equisingular" for ϖ. (cf. Figure 29.)

The meaning of this condition is that the pullbacks of a solution have the same negative pieces of the Birkhoff factorization, independent of the choice of the section and hence of the parameter μ, in our physical interpretation of the geometric data. In the proof of Theorem 1.67 below, we will see much more precisely this relation between equisingular connections and loops in $L(G(\mathbb{C}), \mu)$.

We introduce the following notion of equivalence for connections on P^0.

Definition 1.66. *Two connections ϖ and ϖ' on P^0 are equivalent iff*

$$\varpi' = Dh + h^{-1}\varpi h, \tag{1.302}$$

with h a G-valued $\mathbb{G}_m$-invariant map regular in B.

We again fix a choice of a (non-canonical) regular section $\sigma : \Delta \to B$, with $\sigma(0) = y_0$. For simplicity of notation, we write $\Delta \times \mathbb{G}_m$ for the base space, identified with B through the trivialization given by the section σ.

The main result of this section is the following classification of flat equisingular connections.

Theorem 1.67. *Consider the space $P = B \times G$ as above. There is a bijective correspondence between equivalence classes of flat equisingular G-connections ϖ on P^0 and elements $\beta \in \mathfrak{g}(\mathbb{C})$, with the following properties:*

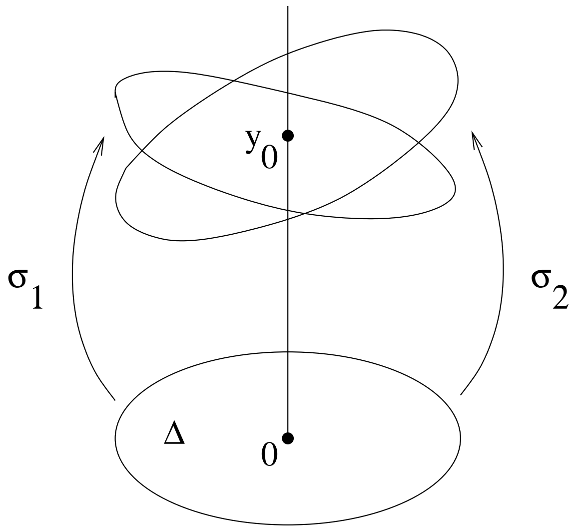


FIGURE 29. Equisingular connections have the property that, when approaching the singular fiber, the type of singularity of $\sigma^*\gamma$, for $D\gamma = \varpi$, does not depend on the section σ along which one restricts γ, but only on the value of the section at $z = 0$.

(1) *Given a flat equisingular connection* ϖ, *there is a unique* $\beta \in \mathfrak{g}(\mathbb{C})$ *such that* $\varpi \sim D\gamma$ *for*

$$\gamma(z, v) = \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\beta)\,\frac{du}{u}}, \tag{1.303}$$

with the integral performed on the straight path $u = tv$, $t \in [0, 1]$.

(2) *This correspondence is independent of the choice of a local regular section* $\sigma : \Delta \to B$ *with* $\sigma(0) = y_0$.

PROOF. (1) We show that, for a given flat equisingular G-connection ϖ, there exists a unique element $\beta \in \mathfrak{g}(\mathbb{C})$ such that ϖ is equivalent, in the sense of Definition 1.66, to the flat equisingular connection $D\gamma$, with $\gamma(z, v)$ as in (1.303).

This will be achieved in three main steps. We first need to prove the vanishing of the monodromies $M_{\Delta^*}(\varpi)$ and $M_{\mathbb{C}^*}(\varpi)$ associated to the two generators of the fundamental group $\pi_1(B)$. This is necessary to ensure the existence of a solution to the equation $D\gamma = \varpi$. The next step is to get around the problem of the choice of the base point. This will be done by showing that an invariant connection automatically extends to the product $\Delta^* \times \mathbb{C}$ while the restriction to $\Delta^* \times \{0\}$ vanishes. Thus one can take arbitrarily a base point in $\Delta^* \times \{0\}$ and get a canonical formula for solutions corresponding to such a specific choice of base point. We then use the equisingularity condition and apply the result of Theorem 1.60 to the restriction of γ to a section of B over Δ.

Using the trivialization given by $\sigma : \Delta \to B$, we can write the connection ϖ on P^0 in terms of $\mathfrak{g}(\mathbb{C})$-valued 1-forms on B^0 as

$$\varpi = A(z, v)\,dz + B(z, v)\,\frac{dv}{v},$$

where $A(z,v)$ and $B(z,v)$ are $\mathfrak{g}$-valued functions and $\frac{dv}{v}$ is the fundamental 1-form of the principal $\mathbb{C}^*$-bundle B.

By the first condition (1.300) of equisingularity, we have

$$\varpi(z,uv) = u^Y(\varpi(z,v)),$$

which shows that the coefficients A and B of ϖ are determined by their restriction to $v=1$. Thus, we can write ϖ as

$$\varpi(z,u) = u^Y(a(z))\,dz + u^Y(b(z))\,\frac{du}{u}, \tag{1.304}$$

for some $a,b \in \mathfrak{g}(K)$.

The flatness condition means that these coefficients satisfy

$$\frac{db}{dz} - Y(a) + [a,b] = 0. \tag{1.305}$$

The positivity of the integral grading Y on $\mathcal{H}$ shows that the connection ϖ extends to a flat connection on the product $\Delta^* \times \mathbb{C}$. Indeed for any element $a \in \mathfrak{g}(K)$ and $X \in \mathcal{H}$ the function

$$u \mapsto \langle u^Y(a), X\rangle$$

is a polynomial $P(u)$ with $P(0)=0$. This shows that the term $u^Y(b)\,\frac{du}{u}$ in (1.304) is regular at $u=0$. Moreover the restriction of the connection ϖ to $\Delta^* \times \{0\}$ is equal to zero, since $u^Y(a)=0$ for $u=0$.

Consider then the monodromy $M_{\{z_0\}\times\mathbb{C}^*}(\varpi)$ for a fixed value $z_0 \in \Delta^*$. This is trivial, because the connection ϖ extends to the simply connected domain $\{z_0\}\times\mathbb{C}$.

Notice that here we are only working with an infinitesimal disk Δ, while the argument given above works for a disk of finite radius, but in fact we can proceed as in Lemma 1.54, writing the group scheme G as an inverse limit of G_i with finite-dimensional $\mathcal{H}(i)$. We can compute the monodromy $M_{\{z_0\}\times\mathbb{C}^*}(\varpi)$ at this finite level G_i, with $z_0 \in \Delta_i^*$, using the argument above. We can then use the fact that the vanishing of the monodromy is a well defined condition in the projective limit.

Consider the monodromy $M_{\Delta^*\times\{u\}}(\varpi)$. Notice that we can choose to compute this monodromy for the value $u=0$, but the restriction of the connection to $\Delta^*\times\{0\}$ is identically equal to 0, so this monodromy also vanishes.

We have shown that the connection ϖ has trivial monodromy along the two generators of $\pi_1(B^0)=\mathbb{Z}^2$. Thus, we know that the equation $D\gamma=\varpi$ admits solutions, and in fact we can write explicitly a solution as in Proposition 1.56, taking as the base point $(z_0,0)\in\Delta^*\times\{0\}$. Namely, we consider a path in $\Delta^*\times\{0\}$ from $(z_0,0)$ to $(z,0)$ and then the straight path (z,tv), $t\in[0,1]$. By Proposition 1.52, this gives a solution of the form

$$\gamma(z,v) = \mathrm{T}e^{\int_0^v u^Y(b(z))\,\frac{du}{u}}\,, \tag{1.306}$$

with $b(z)$ as in (1.304) and with the integral performed on the straight path $u=tv$, for $t\in[0,1]$.

Notice that we have constructed an *invariant* solution of the equation $D\gamma=\varpi$, i.e. one has

$$\gamma(z,wu) = w^Y\gamma(z,u). \tag{1.307}$$

It is not true in general that an invariant connection on an equivariant bundle admits invariant solutions, precisely because of the problem of the choice of an invariant base point. This problem is dealt with above by the choice of a base point in $\Delta^* \times \{0\}$.
Let $\gamma(z) = \gamma(z,v)|_{v=1}$. By construction $\gamma(z)$ satisfies

$$\gamma(z,u) = u^Y \gamma(z). \tag{1.308}$$

Since $\gamma(z,v)$ is a solution of $D\gamma = \varpi$, the loop $\gamma(z)$ also satisfies

$$\gamma(z)^{-1} d\gamma(z) = a(z)\, dz \quad \text{and} \quad \gamma(z)^{-1} Y \gamma(z) = b(z)\,. \tag{1.309}$$

Let us denote by σ_s the section $\sigma_s(z) = (z, e^{sz})$, for $z \in \Delta$. The restriction of (1.308) to the section σ_s can be written in the form

$$\gamma_s(z) = \theta_{sz} \gamma(z). \tag{1.310}$$

The second condition (1.301) of equisingularity for ϖ implies that the negative parts of the Birkhoff factorization $\gamma_s(z) = \gamma_{s^-}(z)^{-1} \gamma_{s^+}(z)$ satisfy

$$\frac{\partial}{\partial s} \gamma_{s^-}(z) = 0. \tag{1.311}$$

It is immediate to compare (1.310) and (1.311) with the conditions (1.240) and (1.241) of Definition 1.45. Thus, in particular, we have obtained that a flat equisingular connection ϖ on P^0 determines a loop $\gamma_s(z)$ in the class $L(G(\mathbb{C}), \mu)$, for $\mu = e^s$.
We can then apply the classification result for loops in $L(G(\mathbb{C}), \mu)$, proved in Theorem 1.60. We obtain that there exists a unique element $\beta \in \mathfrak{g}(\mathbb{C})$ and a regular loop $\gamma_{\text{reg}}(z)$ such that

$$\gamma(z,1) = \mathrm{T}e^{-\frac{1}{z} \int_\infty^0 \theta_{-t}(\beta)\, dt}\, \gamma_{\text{reg}}(z)\,. \tag{1.312}$$

We obtain from (1.312) the expression

$$\gamma(z,v) = v^Y (\mathrm{T}e^{-\frac{1}{z} \int_\infty^0 \theta_{-t}(\beta)\, dt})\, v^Y(\gamma_{\text{reg}}(z)). \tag{1.313}$$

Since v^Y is an automorphism, we also have

$$v^Y (\mathrm{T}e^{-\frac{1}{z} \int_\infty^0 \theta_{-t}(\beta)\, dt}) = \mathrm{T}e^{-\frac{1}{z} \int_0^v u^Y(\beta)\, \frac{du}{u}}, \tag{1.314}$$

with the second integral taken on the straight path $u = tv$, for $t \in [0,1]$. Thus, we can write (1.313) as

$$\gamma(z,v) = \left(\mathrm{T}e^{-\frac{1}{z} \int_0^v u^Y(\beta)\, \frac{du}{u}}\right) v^Y(\gamma_{\text{reg}}(z)). \tag{1.315}$$

Thus, using $h = v^Y(\gamma_{\text{reg}}(z))$ as the regular loop realizing the equivalence (1.302), we obtain, as stated, the equivalence of flat equisingular connections

$$\varpi \sim D(\mathrm{T}e^{-\frac{1}{z} \int_0^v u^Y(\beta)\, \frac{du}{u}}).$$

We now need to understand how the class of the solution (1.313) depends upon $\beta \in \mathfrak{g}$. Consider an equivalence $\varpi' = Dh + h^{-1} \varpi h$, with h a G-valued $\mathbb{G}_m$-invariant map regular in B, as in Definition 1.66. Since h is $\mathbb{G}_m$-invariant one has $h(z,u) = u^Y(h(z,1)$ and since we are dealing with group elements it extends regularly to $h(z,0) = 1$. Thus both the solutions γ_j and the map h are normalized to be 1 on

$\Delta^* \times \{0\}$ and the equivalence generates a relation between invariant solutions of the form

$$\gamma_2(z,u) = \gamma_1(z,u)\,h(z,u), \tag{1.316}$$

with h regular. Thus, for $j = 1, 2$, the loops

$$\gamma_j(z,1) = \mathrm{T}e^{-\frac{1}{z}\int_\infty^0 \theta_{-t}(\beta_j)\,dt}$$

will have the same negative pieces of the Birkhoff factorization, $\gamma_{1^-} = \gamma_{2^-}$. This implies $\beta_1 = \beta_2$, by the equality of the residues at $z = 0$.

Finally, suppose given an element $\beta \in \mathfrak{g}(\mathbb{C})$ and consider the loop

$$\gamma(z,v) = \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\beta)\,\frac{du}{u}}. \tag{1.317}$$

We need to show that $\varpi = D\gamma$ is equisingular. The equivariance condition (1.300) is satisfied by construction since γ is invariant. To see that the second condition (1.301) is also satisfied, we first note that it is enough to check it for one solution γ of the equation $D\gamma = \varpi$. Indeed, any other solution is of the form $\gamma' = g\,\gamma$ with $g \in G(\mathbb{C})$. The relation between the negative parts of the Birkhoff decompositions is the conjugacy by g and this preserves the equisingularity condition. Thus we just need to check the latter on the solution γ of the form (1.317). We let $v(z) \in \mathbb{C}^*$ be a regular function of $z \in \Delta$, with $v(0) = 1$ and consider the section $v(z)\sigma(z)$ instead of the chosen trivialization $\sigma(z)$. The restriction of the solution γ to this section is given by the $G(\mathbb{C})$-valued loop

$$\gamma_v(z) = \mathrm{T}e^{-\frac{1}{z}\int_0^{v(z)} u^Y(\beta)\,\frac{du}{u}}\;. \tag{1.318}$$

Then the Birkhoff factorization $\gamma_v(z) = \gamma_{v^-}(z)^{-1}\gamma_{v^+}(z)$ is given by

$$\gamma_{v^-}(z)^{-1} = \mathrm{T}e^{-\frac{1}{z}\int_0^1 u^Y(\beta)\,\frac{du}{u}}, \quad \text{and} \quad \gamma_{v^+}(z) = \mathrm{T}e^{-\frac{1}{z}\int_1^{v(z)} u^Y(\beta)\,\frac{du}{u}}. \tag{1.319}$$

In fact, we have

$$\gamma_v(z) = \left(\mathrm{T}e^{-\frac{1}{z}\int_0^1 u^Y(\beta)\,\frac{du}{u}}\right)\left(\mathrm{T}e^{-\frac{1}{z}\int_1^{v(z)} u^Y(\beta)\,\frac{du}{u}}\right), \tag{1.320}$$

where the first term in the product is a regular function of z^{-1} and gives a polynomial in z^{-1} when paired with any element of $\mathcal{H}$, while the second term is a regular function of z, as one can see using the Taylor expansion of $v(z)$ at $z = 0$, with $v(0) = 1$.

(2) This last argument, which we used to prove the equisingularity of $D\gamma$, describes explicitly the effect of changing the choice of the trivialization $\sigma : \Delta \to B$ by a factor given by a regular function $v(z) \in \mathbb{C}^*$. Thus, the same argument also gives us as a consequence the independence of the the choice of the trivialization of B.

□

We have the following result, which follows from the previous argument.

Corollary 1.68. *Let ϖ be a flat equisingular connection. Suppose given two such choices of trivialization $\sigma_j : \Delta \to B$, with $\sigma_2 = \alpha\,\sigma_1$, for some $\alpha(z) \in \mathbb{C}^*$ regular at $z = 0$ with $\alpha(0) = 1$. Then the regular values $\gamma_{\mathrm{reg}}(y_0)_j$ of solutions of the differential systems associated to the connections $\sigma_i^*(\varpi)$ are related by*

$$\gamma_{\mathrm{reg}}(y_0)_2 = e^{-s\,\beta}\;\gamma_{\mathrm{reg}}(y_0)_1, \tag{1.321}$$

with

$$s = \frac{d\alpha(z)}{dz}|_{z=0} \,.$$

REMARK 1.69. In the physics context, the relation (1.321) expresses the ambiguity inherent in the renormalization group action as coming from the fact that there is no preferred choice of a local regular section σ of B.

We have obtained, by Theorem 1.67 and Theorem 1.60, an explicit correspondence between the data $\gamma_-(z)$ for loops $\gamma_\mu(z)$ in $L(G(\mathbb{C}),\mu)$ and equivalence classes of flat equisingular connections. In particular, in the physically relevant case where $G = \mathrm{Difg}(\mathcal{T})$, this makes it possible to encode geometrically the data of the counterterms (the divergences) of the renormalizable theory $\mathcal{T}$ as a class of equisingular flat connections over the space P described above.

7.3. Equivariant principal bundles and the group $G^* = G \rtimes \mathbb{G}_m$.

The formalism of the above discussion was based on $\mathfrak{g}(\mathbb{C})$-valued differential forms on the base space B. We now relate more precisely this discussion with connections on $\mathbb{G}_m$-equivariant principal bundles over B. This will involve the affine group scheme $G^* = G \rtimes \mathbb{G}_m$ of Lemma 1.25.

To fix the notation, we recall that, given a principal bundle

$$\pi : P \to B\,, \quad \text{with} \quad P \times H \to P\,, \quad (\xi, h) \mapsto R_h(\xi) = \xi\, h, \tag{1.322}$$

with structure group H acting on the right, a connection is specified by a 1-form α on P with values in the Lie algebra $\mathfrak{h}$ of H and such that

$$\alpha|_{\pi^{-1}(b)} = h^{-1}dh\,, \quad \text{with} \quad R_a^*(\alpha) = (\mathrm{Ad}\, a^{-1})(\alpha), \tag{1.323}$$

where the Maurer-Cartan form $h^{-1}dh$ makes sense in the fibers $\pi^{-1}(b) \sim H$ independently of the choice of a base point. Given a section $\xi : B \to P$ and a connection on P with connection form α, the pullback $\varpi = \nabla(\xi) = \xi^*(\alpha)$ is a 1-form on B with values in the Lie algebra $\mathfrak{h}$. For any smooth map k from B to H, one has

$$\nabla(\xi\, k) = \, k^{-1}\, dk + \, k^{-1}\, \nabla(\xi)\, k. \tag{1.324}$$

In particular, the knowledge of the 1-form $\varpi = \nabla(\xi) = \xi^*(\alpha)$ uniquely determines the connection ∇. Moreover, if we fix the section ξ and the connection ∇, then looking for a flat section $\eta = \xi\, k$, with $\nabla\eta = 0$, is the same thing as trying to find a smooth map k from B to H such that

$$k^{-1}\, dk + \, k^{-1}\, \nabla(\xi)\, k = \, 0. \tag{1.325}$$

Equivalently, in terms of $h = k^{-1}$ and $\varpi = \nabla(\xi)$, it means solving the equation

$$h^{-1}\, dh = \, \varpi. \tag{1.326}$$

Let now $H_1 \subset H$ be a closed subgroup that acts on the base B. We consider the trivial H-principal bundle

$$P = B \times H\,, \quad \text{with} \quad R_a(b,h) = \, (b, ha)\,, \quad \forall b \in B\,, \quad h \in H\,, \quad a \in H. \tag{1.327}$$

PROPOSITION 1.70. *1) The action*

$$h_1\,(b,h) = (h_1\,b, h_1\,h) \tag{1.328}$$

of H_1 on P turns this H-principal bundle into an H_1-equivariant H-principal bundle.
2) Let ∇ be a connection on P, and $\varpi = \nabla(\xi)$, where ξ is the section $\xi(b) = (b,1)$. Then, for any $h_1 \in H_1$, the pull-back ∇' of the connection ∇ by the action of h_1 on P is the connection given by

$$\varpi' = \nabla'(\xi) = h_1^{-1}\,h_1^*(\varpi)\,h_1\,. \tag{1.329}$$

PROOF. 1) One needs to check that the projection π is equivariant, which is clear, and that the left action of h_1 commutes with the right action of the group H, which is also clear.
2) Let α be the connection 1-form of the connection ∇. Let $h_1 \in H_1$. By construction the connection 1-form α' of the pullback ∇' is the pullback of α by the left multiplication L of (1.328). Thus, we obtain

$$\nabla'(\xi) = \xi^*(\alpha') = \xi^*(L^*(\alpha)) = (L\circ\xi)^*(\alpha).$$

One has

$$L\circ\xi(b) = (h_1\,b, h_1) = R_{h_1}\,\xi(h_1\,b)\,,\quad \forall b\in B.$$

By (1.323) $R_{h_1}^*(\alpha) = (\mathrm{Ad}\,h_1^{-1})(\alpha)$, which gives the required equality (1.329). ☐

PROPOSITION 1.71. *1) A connection ∇ on P is H_1-invariant iff $\varpi = \nabla\xi$ satisfies the condition*

$$h_1^*(\varpi) = h_1\,\varpi\,h_1^{-1}\,,\quad \forall h_1 \in H_1\,. \tag{1.330}$$

2) Let γ be a smooth map from B to H such that

$$\gamma(h_1\,b) = h_1\,\gamma(b)\,h_1^{-1}\,,\quad \forall h_1 \in H_1\,. \tag{1.331}$$

then the gauge equivalence

$$\varpi \mapsto \varpi' = \gamma^{-1}d\gamma + \gamma^{-1}\,\varpi\,\gamma \tag{1.332}$$

preserves the H_1-invariance of connections.

PROOF. 1) This follows from (1.329).
2) Let us assume that ϖ fulfills (1.330). We show that ϖ' also does. One has

$$h_1^*(\gamma^{-1}\,\varpi\,\gamma) = h_1^*(\gamma)^{-1}\,h_1^*(\varpi)\,h_1^*(\gamma) = h_1\,\gamma^{-1}\,\varpi\,\gamma\,h_1^{-1}$$

and since $dh_1 = 0$ one has

$$h_1^*(\gamma^{-1}\,d\gamma) = h_1\,\gamma^{-1}\,d\gamma\,h_1^{-1},$$

so that the required invariance follows. ☐

We now specialize Propositions 1.70 and 1.71 to the subgroup

$$H_1 = \mathbb{G}_m \subset H = G^* = G \rtimes \mathbb{G}_m$$

and to connections with values in the Lie subalgebra

$$\mathfrak{g} = \mathrm{Lie}(G) \subset \mathfrak{g}^* = \mathrm{Lie}(G^*) \tag{1.333}$$

The adjoint action of the subgroup $H_1 = \mathbb{G}_m$ gives the grading u^Y on $\mathfrak{g} = \mathrm{Lie}(G)$.
A comment about notation: since we are interested in the relative situation of the inclusion $B^0 \subset B$, we adopt the notation X^0 to denote the restriction to $B^0 =$

$B \setminus V \subset B$ of an object X (bundle, connection, section, etc) defined on B. In particular we let $\tilde{P} = B \times G^*$ be the trivial principal G^*-bundle over the base space B, and $\tilde{P}^0$ its restriction to $B^0 = B \setminus V \subset B$,

$$\tilde{P}^0 = B^0 \times G^* \tag{1.334}$$

We have the following result.

PROPOSITION 1.72. *Let $\mathbb{G}_m \to B \xrightarrow{\pi} \Delta$ denote the $\mathbb{G}_m(\mathbb{C}) = \mathbb{C}^*$-bundle over the infinitesimal disk Δ of* (1.298). *We let $\tilde{P} = B \times G^*$ be the trivial principal G^*-bundle over the base space B.*
1) The action

$$u\,(b,k) = \,(ub, uk) \tag{1.335}$$

of $\mathbb{G}_m$ on $\tilde{P}$ turns this G^*-principal bundle into a $\mathbb{G}_m$-equivariant G^*-principal bundle.
2) Let $\xi(b) = (b,1)$. Then the map

$$\nabla \mapsto \nabla(\xi) \tag{1.336}$$

is an isomorphism between $\mathfrak{g}$-valued $\mathbb{G}_m$-invariant connections on $\tilde{P}^0$ and $\mathfrak{g}$-valued connections on B^0 fulfilling (1.300).
3) Let γ be a regular map from B to G such that

$$\gamma(ub) = \,u\,\gamma(b)\,u^{-1}\,, \quad \forall u \in \mathbb{G}_m\,. \tag{1.337}$$

Then the transformation L_γ, given by $L_\gamma(b,\,k) = (b, \gamma(b)k)$, is an automorphism of the $\mathbb{G}_m$-equivariant G^*-principal bundle $\tilde{P}$.
4) The gauge equivalence of connections under the automorphisms L_γ corresponds under the isomorphism (1.336) *to the equivalence of Definition* 1.66.

PROOF. 1) The construction of the $\mathbb{G}_m$-equivariant G^*-principal bundle follows from Proposition 1.70.
2) The invariance condition on connections given by (1.300) is identical to (1.330).
3) By construction L_γ is an automorphism of G^*-principal bundle. Condition (1.337) means that it commutes with the action of $\mathbb{G}_m$.
4) Let α be the connection form of the first connection. The connection form of the transformed connection is $L_\gamma^*(\alpha)$, so that the actions on sections are related by $\nabla' = \nabla \circ L_\gamma$. Then $\varpi' = \nabla'(\xi) = \nabla(\tilde{\gamma})$ where $\tilde{\gamma}(b) = (b,\, \gamma(b))$. Thus, we write

$$\varpi' = \,\gamma^{-1}\,d\gamma + \,\gamma^{-1}\,\varpi\,\gamma.$$

The invariance condition on gauge equivalences given by (1.337) is identical to the one used for the equivalence of equisingular connections in Definition 1.66. ☐

It is important to understand geometrically the meaning of the restriction to $\mathfrak{g}$-valued connections in part 2) of Proposition 1.71 and to maps to G (instead of G^*) in part 3) of the same proposition. In both cases one needs to express the triviality of the image under the canonical morphism

$$\epsilon : G^* = G \rtimes \mathbb{G}_m \to \mathbb{G}_m. \tag{1.338}$$

We let $\epsilon(\tilde{P})$ be the $\mathbb{G}_m$-equivariant $\mathbb{G}_m$-principal bundle obtained from $\tilde{P}$ and the homomorphism ϵ. We denote by $\tilde{\epsilon}$ the bundle map $\tilde{\epsilon} : \tilde{P} \to \epsilon(\tilde{P})$.

LEMMA 1.73. *1) A connection ∇ on $\tilde{P}^0$ is $\mathfrak{g}$-valued iff it is compatible through $\tilde{\epsilon}$ with the trivial connection on $\epsilon(\tilde{P})$, i.e. iff $\epsilon(\nabla) = d$.*
2) Suppose given an invariant flat connection ∇ on $\tilde{P}^0$ with $\epsilon(\nabla) = d$ as above. For any section η with $\nabla \eta = 0$, consider the unique bundle isomorphism between the restrictions of $\tilde{P}^0$ to sections $\sigma : \Delta \to B$ with $\sigma(0) = y_0$, determined by the trivialisation η. For such an invariant flat connection ∇, consider the corresponding $\nabla(\xi)$, defined as in (1.336)*. This $\nabla(\xi)$ is equisingular iff the bundle isomorphism defined by η is regular on Δ.*
3) An invariant bundle automorphism of the $\mathbb{G}_m$-equivariant G^*-principal bundle $\tilde{P}$ comes from a map γ as in (1.337) *iff it induces the identity automorphism on $\epsilon(\tilde{P})$.*

PROOF. 1) The compatibility with the trivial connection means that the image by $\tilde{\epsilon}$ of horizontal vectors is horizontal. Let $\xi(b) = (b, 1)$ and $\varpi = \nabla(\xi)$. The compatibility holds iff $\epsilon(\varpi(X)) = 0$ for every tangent vector, i.e. iff ϖ, which is a priori $\mathfrak{g}^*$-valued, is in fact $\mathfrak{g}$-valued.
2) Let $\varpi = \nabla(\xi)$. A solution $\gamma(z, v) \in G$ of the equation $D\gamma = \varpi$ gives a flat section $\eta(b) = (b, \gamma^{-1}(b))$ of $\tilde{P}^0$. Let $\sigma_j : \Delta \to B$ with $\sigma_j(0) = y_0$. The bundle isomorphism between the $\sigma_j^* \tilde{P}^0$, associated to the trivialisation η, is given by left multiplication by a map $z \mapsto k(z) \in G^*$, such that

$$k(z)\, \gamma^{-1}(\sigma_1(z)) = \gamma^{-1}(\sigma_2(z))\,, \quad \forall z \in \Delta^*.$$

Thus, one has $k = \sigma_2^*(\gamma)^{-1}\, \sigma_1^*(\gamma)$ and k is regular iff the $\sigma_j^*(\gamma)$ have the same singularity.
This shows that, if a flat section η for ∇ fulfills the condition of statement 2) of the lemma, then $\varpi = \nabla(\xi)$ is equisingular. Conversely, if $\varpi = \nabla(\xi)$ is equisingular, then we get a flat section η of $\tilde{P}^0$ which fulfills the condition of 2) of the lemma.
In fact, any other flat section is of the form $\eta' = \eta g$ for some $g \in G^*(\mathbb{C})$, and this does not alter the regularity of the bundle isomorphism associated to the trivialization.
3) A bundle automorphism is given by the left multiplication by a map γ from B to G^*. It is invariant iff (1.337) holds, and it induces the identity automorphism iff $\epsilon \circ \gamma = 1$. This means that γ takes values in $G = \mathrm{Ker}\, \epsilon$. □

DEFINITION 1.74. *We say that ∇ is equisingular when it fulfills the hypothesis of 2) of Lemma* 1.73*.*

Thus, we have translated into this more geometric language the notion of equisingular connection of Definition 1.65.

We remarked in the proof of Theorem 1.67 that the flat invariant connections extend automatically to the "compactification" of the base B obtained by passing from $\mathbb{C}^*$ to $\mathbb{C}$. We now state this result in more geometric terms. We define the "compactification" of the base B as

$$\underline{B} = B \times_{\mathbb{G}_m} \mathbb{G}_a. \tag{1.339}$$

This is still a $\mathbb{G}_m$-space, since one has a natural action of $\mathbb{G}_m$ on the right, but it is no longer a principal $\mathbb{G}_m$-bundle, since this action admits a whole set of fixed points given by

$$B_{\mathrm{fix}} = B \times_{\mathbb{G}_m} \{0\} \subset \underline{B} = B \times_{\mathbb{G}_m} \mathbb{G}_a. \tag{1.340}$$

As above, we use the notation X^0 to indicate the restriction of a bundle X to B^0.

LEMMA 1.75. *1) The bundle $\tilde{P}$ extends canonically to a $\mathbb{G}_m$-equivariant G^*-principal bundle $\underline{\tilde{P}}$ on $\underline{B}$.*
2) Any invariant connection ∇ on $\tilde{P}^0$ with $\epsilon(\nabla) = d$ extends canonically to an invariant connection $\underline{\nabla}$ on $\underline{\tilde{P}}^0$. The restriction of $\underline{\nabla}$ to B_{fix} is the trivial connection d.
3) An invariant bundle automorphism γ of the $\mathbb{G}_m$-equivariant G^*-principal bundle $\tilde{P}$ with $\epsilon(\gamma) = 1$ extends to an invariant bundle automorphism of $\underline{\tilde{P}}$ whose restriction to B_{fix} is the identity.

PROOF. 1) Let $\underline{\tilde{P}}$ be the trivial G^*-principal bundle $\underline{\tilde{P}} = \underline{B} \times G^*$. Since the action of $\mathbb{G}_m$ on B extends canonically to an action on $\underline{B}$ we can apply Proposition 1.70 1) and get the required $\mathbb{G}_m$-equivariant G^*-principal bundle.
2) Let us show that the differential form $\varpi = \nabla \xi$ on B^0, where ξ is the section $\xi(b) = (b, 1)$, extends to $\underline{B}^0$ by continuity. This extension then uniquely determines the extension of ∇ as a connection on $\underline{\tilde{P}}^0$. We use a (non-canonical) trivialization of B as a $\mathbb{G}_m$-principal bundle, and write $B = \Delta \times \mathbb{G}_m$. The condition $\epsilon(\nabla) = d$ shows that the form ϖ takes values in the Lie subalgebra $\mathfrak{g} \subset \mathfrak{g}*$, where $\mathfrak{g}* = \text{Lie}(G^*)$. Thus, it is of the form

$$\varpi = a(z,v)\, dz \,+\, b(z,v)\, \frac{dv}{v},$$

where $a(z,v)$ and $b(z,v)$ are $\mathfrak{g}$-valued functions and $\frac{dv}{v}$ is the fundamental 1-form of the principal $\mathbb{C}^*$-bundle B. The invariance condition means that

$$a(z,uv) = u^Y\, a(z,v)\,, \quad b(z,uv) = u^Y\, b(z,v)\,, \quad \forall u \in \mathbb{G}_m(\mathbb{C}).$$

Thus, the strict positivity of the grading Y on $\mathfrak{g}$ shows that $a(z,v)\,dz$ extends by 0 for $v = 0$ and that $u^Y\, b(z,1)\frac{1}{u}$ extends by continuity to $\underline{B}^0$ (with a not necessarily zero value at $u = 0$) so that $b(z,v)\, \frac{dv}{v}$ extends to $\underline{B}^0$.
3) The invariant bundle automorphism is given by left multiplication L_γ by $\gamma(z,v) \in G$ fulfilling

$$\gamma(z,uv) = u^Y \gamma(z,v)\,, \quad \forall u \in \mathbb{G}_m(\mathbb{C}).$$

Since the grading is strictly positive on the kernel of the augmentation ideal of the Hopf algebra of the affine group scheme G, one gets that $u^Y\, \gamma(z,1)$ extends by continuity to $\gamma(z,0) = 1$. $\square$

REMARK 1.76. In the physics interpretation of Remark 1.64, the fiber

$$\pi^{-1}(\{z\}) \subset B$$

over z is the set of possible values of $\mu^z\, \hbar$. Thus the compactification $\underline{B}$ corresponds then to adding the classical limit $\hbar = 0$.

7.4. Tannakian categories and affine group schemes.

In the previous sections we have translated the data of perturbative renormalization from loops to equisingular flat connections. This makes it possible to use the Riemann–Hilbert correspondence and classify the equisingular flat connections in terms of representation theoretic data.

We begin by recalling some general facts about the Tannakian formalism. This gives a very general setting in which a given set of data with suitable properties (a neutral Tannakian category) can be shown to be equivalent to the category finite-dimensional linear representations of an affine group scheme.

An abelian category is a category to which the usual tools of homological algebra apply. This is made precise by the following definition (cf. e.g. [**139**], §2.2).

DEFINITION 1.77. *A category $\mathcal{C}$ is an abelian category if the following axioms are satisfied.*

- *For any $X, Y \in \mathrm{Obj}(\mathcal{C})$, the set $\mathrm{Hom}_{\mathcal{C}}(X, Y)$ is an abelian group, with respect to which the composition of morphisms is bi-additive.*
- *There is an object $0 \in \mathrm{Obj}(\mathcal{C})$ such that $\mathrm{Hom}_{\mathcal{C}}(0, 0)$ is the trivial group.*
- *There are finite products and coproducts, namely, for each $X, X' \in \mathrm{Obj}(\mathcal{C})$ there exists a $Y \in \mathrm{Obj}(\mathcal{C})$ and morphisms*
$$X \xrightarrow{f_1} Y \xleftarrow{f_2} X' \quad \text{and} \quad X \xleftarrow{h_1} Y \xrightarrow{h_2} X',$$
with $h_1 f_1 = 1_X$, $h_2 f_2 = 1_{X'}$, $h_2 f_1 = 0 = h_1 f_2$, $f_1 h_2 + f_2 h_1 = 1_Y$.
- *For any $X, Y \in \mathrm{Obj}(\mathcal{C})$, every morphism $f : X \to Y$ has a canonical decomposition*
$$K \xrightarrow{k} X \xrightarrow{i} I \xrightarrow{j} Y \xrightarrow{c} K',$$
where $j \circ i = f$, with $K = \mathrm{Ker}(f)$, $K' = \mathrm{Coker}(f)$, and $I = \mathrm{Ker}(k) = \mathrm{Coker}(c)$.

We need to consider more structure, in order to arrive at something that has the right properties to compare with the category of representations of an affine group scheme.

DEFINITION 1.78. *A category $\mathcal{C}$ is k-linear for a field k if, for any $X, Y \in \mathrm{Obj}(\mathcal{C})$, the set $\mathrm{Hom}_{\mathcal{C}}(X, Y)$ is a k-vector space. A tensor category over k is a k-linear category $\mathcal{C}$, endowed with a bi-functor $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ and a distinguished object $1 \in \mathrm{Obj}(\mathcal{C})$ satisfying the following property:*

- *There are functorial isomorphisms*
$$a_{X,Y,Z} : X \otimes (Y \otimes Z) \to (X \otimes Y) \otimes Z$$
$$c_{X,Y} : X \otimes Y \to Y \otimes X$$
$$l_X : X \otimes 1 \to X \quad \text{and} \quad r_X : 1 \otimes X \to X.$$
One assumes in the commutativity property that $c_{Y,X} = c_{X,Y}^{-1}$.

A rigid tensor category is a tensor category $\mathcal{C}$ over k with a duality $\vee : \mathcal{C} \to \mathcal{C}^{op}$, satisfying the following properties:

- *For any $X \in \mathrm{Obj}(\mathcal{C})$ the functor $- \otimes X^\vee$ is left adjoint to $- \otimes X$ and the functor $X^\vee \otimes -$ is right adjoint to $X \otimes -$.*
- *There is an evaluation morphism $\epsilon : X \otimes X^\vee \to 1$ and a unit morphism $\delta : 1 \to X^\vee \otimes X$ satisfying $(\epsilon \otimes 1) \circ (1 \otimes \delta) = 1_X$ and $(1 \otimes \epsilon) \circ (\delta \otimes 1) = 1_{X^\vee}$.*

One assumes that $\mathrm{End}(1) \cong k$.

REMARK 1.79. *There is a well-defined notion of dimension for objects in a rigid tensor category, with values in $k \cong \mathrm{End}(1)$, given by $\dim(X) := \mathrm{Tr}(1_X) = \epsilon \circ c_{X^\vee,X} \circ \delta$.*

We follow here the terminology of [**106**] p.165. Elsewhere in the literature the same quantity $\mathrm{Tr}(1_X)$ is referred to as the Euler characteristic. The reason for considering it a "dimension" lies in the result recalled in Remark 1.82 below. For the relation in the motivic context of this notion to the usual Euler characteristic and its effect on the Tannakian property of the category of pure motives see [**2**].

Recall also the following properties of functors.

DEFINITION 1.80. *A functor $\omega : \mathcal{C} \to \mathcal{C}'$ is faithful if, for all $X, Y \in \mathrm{Obj}(\mathcal{C})$, the mapping*

$$\omega : \mathrm{Hom}_{\mathcal{C}}(X, Y) \to \mathrm{Hom}_{\mathcal{C}'}(\omega(X), \omega(Y)) \tag{1.341}$$

is injective. If $\mathcal{C}$ and $\mathcal{C}'$ are k-linear categories, a functor ω is additive if (1.341) *is a k-linear map. An additive functor ω is exact if, for any exact sequence $0 \to X \to Y \to Z \to 0$ in $\mathcal{C}$, the corresponding sequence $0 \to \omega(X) \to \omega(Y) \to \omega(Z) \to 0$ in $\mathcal{C}'$ is also exact. A functor $\omega : \mathcal{C} \to \mathcal{C}'$ between k-linear tensor categories is a tensor functor if there are functorial isomorphisms $\tau_1 : \omega(1) \to 1$ and*

$$\tau_{X,Y} : \omega(X \otimes Y) \to \omega(X) \otimes \omega(Y).$$

We further enrich the structure of a rigid tensor category by mapping it to a category of vector spaces through a fiber functor. This is the extra piece of structure that will make it possible to recover the affine group scheme from the category.

DEFINITION 1.81. *Let $\mathcal{C}$ be a k-linear rigid abelian tensor category. Let Vect_K denote the category of finite-dimensional vector spaces over a field K. A fiber functor is an exact faithful tensor functor $\omega : \mathcal{C} \to \mathrm{Vect}_K$. A k-linear rigid abelian tensor category $\mathcal{C}$ is a Tannakian category if it admits a fiber functor $\omega : \mathcal{C} \to \mathrm{Vect}_K$, with K an extension of the field k. The category $\mathcal{C}$ is a neutral Tannakian category if $K = k$.*

REMARK 1.82. *Deligne showed in* [**106**] *that a rigid abelian tensor category $\mathcal{C}$ over a field k of characteristic zero is a Tannakian category if and only if the dimensions $\dim(X) \in \mathrm{End}(1) = k$ are non-negative integers for all objects $X \in \mathrm{Obj}(\mathcal{C})$ (cf. Remark* 1.79*).*

The main example of a neutral Tannakian category is the category Rep_G of finite-dimensional linear representations of an affine group scheme G over k, with fiber functor the forgetful functor to vector spaces. The point of the Tannakian formalism is that all neutral Tannakian categories are of the form Rep_G.

Suppose given a neutral Tannakian category $\mathcal{C}$, with fiber functor $\omega : \mathcal{C} \to \mathrm{Vect}_k$. To see that there is an associated affine group scheme, consider first the group $G(k)$ of automorphisms of ω, i.e. invertible natural transformations that are compatible with all the structure on the category $\mathcal{C}$. If K is an extension of k, then there is an induced functor $\omega_K : \mathcal{C} \to \mathrm{Vect}_K$ by $\omega_K(X) = \omega(X) \otimes_k K$ and one can associate to it the group $G(K)$. Similarly, one can make sense of $G(A)$ for A a unital commutative k-algebra and check that one obtains in this way a covariant representable functor to groups $G : \mathcal{A}_k \to \mathcal{G}$, i.e. an affine group scheme.

We have the following result ([**259**], [**114**]).

THEOREM 1.83. *Let $\mathcal{C}$ be a neutral Tannakian category, with fiber functor $\omega : \mathcal{C} \to \mathrm{Vect}_k$. Let $G = \underline{\mathrm{Aut}}^{\otimes}(\omega)$ be the affine group scheme of automorphisms of the fiber functor. Then ω induces an equivalence of rigid tensor categories*

$$\omega : \mathcal{C} \to \mathrm{Rep}_G.$$

The affine group scheme G is determined by the pair $(\mathcal{C}, \omega)$, but when no confusion can arise as to the fiber functor ω, we simply refer to G as the Galois group of $\mathcal{C}$.

REMARK 1.84. *If $\mathcal{C}' \subset \mathcal{C}$ is a subcategory of a neutral Tannakian category $\mathcal{C}$, which inherits from $\mathcal{C}$ the structure of neutral Tannakian category, then there is a corresponding homomorphism of group schemes $G \to G'$, where $\mathcal{C} \cong \mathrm{Rep}_G$ and $\mathcal{C}' = \mathrm{Rep}_{G'}$.*

The properties of the affine group scheme G reflect the properties of the neutral Tannakian category $\mathcal{C}$. We mention one property that will play a role in §8 below. A k-linear abelian category $\mathcal{C}$ is semi-simple if there exists $A \subset \mathrm{Obj}(\mathcal{C})$ such that all objects X in A are simple (namely $\mathrm{Hom}(X, X) \simeq k$), with $\mathrm{Hom}(X, Y) = 0$ for $X \neq Y$ in A, and such that every object of $\mathcal{C}$ is isomorphic to a direct sum of objects in A. A linear algebraic group over a field of characteristic zero is reductive if every finite-dimensional linear representation is a direct sum of irreducibles. An affine group scheme is pro-reductive if it is a projective limit of reductive linear algebraic groups. These properties are related as follows.

REMARK 1.85. *Let G be the affine group scheme $G = \underline{\mathrm{Aut}}^{\otimes}(\omega)$ of a neutral Tannakian category $\mathcal{C}$ with fiber functor ω. Then G is pro-reductive if and only if the category is semi-simple.*

We look at a concrete case, to illustrate the result of Theorem 1.83. Consider the category Rep_H of finite dimensional complex linear representations of a group H. This is a neutral Tannakian category with fiber functor the forgetful functor to $\mathrm{Vect}_{\mathbb{C}}$. Theorem 1.83 then shows that there exists an affine group scheme G, dual to a commutative Hopf algebra $\mathcal{H}$ over $\mathbb{C}$, such that the fiber functor gives an equivalence of categories $\mathrm{Rep}_H \cong \mathrm{Rep}_G$. This G is called the *algebraic hull* of H and can be quite non-trivial even in very simple cases. For instance, consider the following example (cf. [**283**]).

EXAMPLE 1.86. *Consider the group $H = \mathbb{Z}$. In this case $\mathrm{Rep}_H \cong \mathrm{Rep}_G$, where G is the affine group scheme dual to the Hopf algebra $\mathcal{H} = \mathbb{C}[e(q), t]$, for $q \in \mathbb{C}/\mathbb{Z}$, with the relations $e(q_1 + q_2) = e(q_1)e(q_2)$ and the coproduct $\Delta(e(q)) = e(q) \otimes e(q)$ and $\Delta(t) = t \otimes 1 + 1 \otimes t$.*

The Tannakian formalism was initially proposed by Grothendieck as a "linear version" of the theory of fundamental groups. The main idea is that the group of symmetries (fundamental group in the case of covering spaces, Galois group in the case of algebraic equations) always arises as the group of automorphisms of a fiber functor. In the theory of (profinite) fundamental groups, one considers a functor ω from a certain category of finite étale covers to the category of finite sets. The profinite group $G = \mathrm{Aut}(\omega)$ of automorphisms of this functor determines an equivalence between the category of covers and that of finite G-sets. (The reader interested in seeing more about the case of fundamental groups should look at SGA1 [**154**].)

The Tannakian formalism was then developed by Saavedra [**259**] and Deligne–Milne [**114**] (cf. also the more recent [**106**]). The fiber functor ω with values in vector spaces can be thought of as a "linear version" of the case with values in finite sets, and the affine group scheme $G = \underline{\mathrm{Aut}}^{\otimes}(\omega)$ plays the role of the Galois group (or the fundamental group) in this setting.

In a very different context, a philosophy very similar to the Tannakian formalism can be found in results by Doplicher and Roberts [**124**], obtained with operator algebra techniques in the context of algebraic quantum field theory. There one recovers a global compact gauge group G from its category of finite-dimensional continuous unitary representations. This category is characterized as a monoidal C^*-category whose objects are endomorphisms of certain unital C^*-algebras and whose morphisms are intertwining operators between these endomorphisms. A characterization in terms of integer dimensions similar to the one found by Deligne for Tannakian categories also holds in this context (cf. Remark 1.82).

7.5. Differential Galois theory and the local Riemann–Hilbert correspondence.

For our purposes, we are interested in the Tannakian formalism applied to categories of differential systems, where it has a fundamental role in the context of differential Galois theory and the (local) Riemann–Hilbert correspondence (cf. [**222**], [**284**]).

We recall briefly, in the rest of this section, some general facts about differential Galois theory and the Riemann–Hilbert correspondence, before analyzing the case of equisingular connections in the next section.

Differential Galois theory was discovered by Picard and Vessiot at the beginning of the twentieth century, but it remained for a long time a quite intractable problem to compute the differential Galois group of a given differential equation. All results were confined to the regular singular case, in which it is the Zariski closure of the representation of monodromy, by an old result of Schlesinger. This situation has changed drastically since the work of Deligne, Malgrange, Ramis and Ecalle. In short, Martinet and Ramis discovered a natural generalization of the notion of monodromy that plays in the local analysis of irregular singularities the same role as monodromy does in the regular case.
The fundamental result is that an irregular singular differential equation generates a representation of the wild fundamental group,

$$\pi_1^{\mathrm{wild}} \to \mathrm{GL}_n(\mathbb{C})$$

and this representation classifies the equation up to gauge equivalence. Moreover the differential Galois group of the differential equation is the Zariski closure of the image of π_1^{wild} in $\mathrm{GL}_n(\mathbb{C})$. This holds both in the local and the global case and extends to the irregular case the classical (i.e. regular-singular) Riemann-Hilbert correspondence.

Regular linear differential equation D modulo gauge equivalence	Representation of π_1 in $\mathrm{GL}_n(\mathbb{C})$
$\mathrm{Gal}_K(D)$	Zariski closure of $\mathrm{Im}\,\pi_1$
Linear differential equation D modulo gauge equivalence	Representation of π_1^{wild} in $\mathrm{GL}_n(\mathbb{C})$
$\mathrm{Gal}_K(D)$	Zariski closure of $\mathrm{Im}\,\pi_1^{\mathrm{wild}}$

In the usual setting of differential Galois theory (see e.g. [**284**]), one considers a differential field (K, δ) with field of constants $\mathbb{C}$. The two main cases for the local theory are the formal theory where $K = \mathbb{C}((z))$ and the non-formal theory where $K = \mathbb{C}(\{z\})$.

A system of ordinary differential equations of the form

$$\delta(u) = Au, \tag{1.342}$$

with $A = (a_{ij})$ an $n \times n$ matrix, determines a differential module with $V = K^n$ and the connection ∇ given by $\nabla(e_i) = -\sum a_{ji} e_j$, with the e_i a basis of V over K. The equation (1.342) is then equivalent to the condition $\nabla(\sum_i u_i e_i) = 0$. A linear differential equation

$$Df = 0\,, \qquad \text{with} \quad D = \sum_{j=0}^{n} a_j\, \delta^j \tag{1.343}$$

of order n determines a system, using the companion matrix of the polynomial expression of D.

In general, an equation of the form (1.342) will have at most n solutions in K that are linearly independent over the field of constants, but in general it might not have a full set of solutions in K. In that case, one can consider the Picard–Vessiot extension of K determined by (1.342). This is the smallest differential field extension E of K (the derivation of E restricts to that of K), with the same field of constants, such that (1.342) has a set of n independent solutions in E. The fact that the field of constants is algebraically closed ensures the existence of Picard–Vessiot extensions. The differential Galois group of the equation (1.342) is the group $G(E/K)$ of differential automorphisms of the Picard–Vessiot extension E of (1.342) that fix K. Its action on a system of n independent solutions in E determines a natural faithful representation of $G(E/K)$ in $\mathrm{GL}_n(\mathbb{C})$ and its image is an algebraic subgroup of $\mathrm{GL}_n(\mathbb{C})$.

To understand the meaning of the differential Galois group let us mention two very simple examples and a general result on the *solvability* of equations.

EXAMPLE 1.87. *Let $a \in K$ be an element which does not belong to the image of δ. Then adjoining the primitive of a, that is, a solution of the inhomogeneous equation $\delta(f) = a$, can be described through the system given by the matrix*

$$A = \begin{bmatrix} 0 & -a \\ 0 & 0 \end{bmatrix}.$$

The Picard-Vessiot extension E is given as the differential field $E = K\langle f\rangle$ obtained by adjoining a formal primitive f of a to K. Since $a \notin \operatorname{Im}\delta$ one checks that f is transcendental over K. The differential Galois group is the group $G(E/K) = \mathbb{G}_a(\mathbb{C})$ acting by $f \mapsto f + c$.

In general, it is quite useful, in order to determine the Picard–Vessiot extension E, to know that it is obtained as the field of quotients of the Picard–Vessiot ring R of the differential system (1.342). This is characterized as follows.

- R is a differential algebra;
- R has no non-trivial differential ideal;
- There exists $F \in \mathrm{GL}_n(R)$, such that $F' + A\,F = 0$;
- R is generated by the coefficients of F.

EXAMPLE 1.88. *For given $a \in K$, consider the equation*

$$\delta(u) = au.$$

Let $R = K[u, u^{-1}]$ be obtained by adjoining to K a formal solution and its inverse. If a is such that R has no non-trivial differential ideal, then R is the Picard–Vessiot ring, and the Picard–Vessiot extension E is the field of quotients. In that case the differential Galois group is the group $G(E/K) = \mathbb{G}_m(\mathbb{C})$ acting by $u \mapsto \lambda u$. In general, the differential Galois group is a subgroup of $\mathbb{G}_m(\mathbb{C})$.

The significance of the differential Galois group is that it makes it possible to recognize whether a given equation is solvable by the elementary steps described in Examples 1.87 and 1.88. One has the following general result:

THEOREM 1.89. *The following conditions are equivalent for a linear differential equation D.*
(i) The equation is solvable by repeated applications of the following steps:

- *Adjunction of primitives (example* 1.87*);*
- *Adjunction of exponentials of primitives (example* 1.88*);*
- *Finite algebraic extensions.*

(ii) The connected component (for the Zariski topology) of the identity in the differential Galois group is solvable.

It is quite difficult to compute a differential Galois group in general, but since it is the Zariski closure of the image of the wild fundamental group, the key tool is the Riemann–Hilbert correspondence.

In order to understand the Riemann–Hilbert correspondence in the general local case, we now come to the Tannakian formalism. We first introduce the category of differential modules.

DEFINITION 1.90. *Let (K,δ) be a differential field with $\mathbb{C}$ as field of constants. The category $\mathcal{D}_K$ of differential modules over K has as objects pairs (V,∇) of a vector space $V \in \mathrm{Obj}(\mathcal{V}_K)$ and a connection, namely a $\mathbb{C}$-linear map $\nabla : V \to V$ satisfying $\nabla(fv) = \delta(f)v + f\nabla(v)$, for all $f \in K$ and all $v \in V$. A morphism between objects (V_1,∇_1) and (V_2,∇_2) is a K-linear map $T : V_1 \to V_2$ satisfying $\nabla_2 \circ T = T \circ \nabla_1$.*

The category $\mathcal{D}_K$ of differential modules over K of Definition 1.90 is a $\mathbb{C}$-linear rigid tensor category, with the tensor product

$$(V_1,\nabla_1) \otimes (V_2,\nabla_2) = (V_1 \otimes V_2, \nabla_1 \otimes 1 + 1 \otimes \nabla_2)$$

and with $(V,\nabla)^\vee$ the K-linear dual of V with the induced connection. One can check that the functor to vector spaces over $\mathbb{C}$ that assigns to a module (V,∇) its solution space $\mathrm{Ker}\nabla$ is a fiber functor. Thus, Theorem 1.83 shows that there is an affine group scheme G and an equivalence of categories $\mathcal{D}_K \cong \mathrm{Rep}_G$.

In the formal case $K = \mathbb{C}((z))$, the affine group scheme G is a semidirect product $G = \mathcal{T} \rtimes \bar{\mathbb{Z}}$ of the *Ramis exponential torus* $\mathcal{T}$ by the algebraic hull of $\mathbb{Z}$. The Ramis exponential torus corresponds to the symmetries which multiply by a non-zero scalar

$$e^{P(z^{-1/\nu})} \mapsto \lambda_P \, e^{P(z^{-1/\nu})}, \quad \text{with} \quad \lambda_{P_1+P_2} = \lambda_{P_1} \lambda_{P_2}$$

the formal exponentials of a polynomial in $z^{-1/\nu}$, where ν is a ramification index. This torus is then of the form $\mathcal{T} = \mathrm{Hom}(\mathcal{B},\mathbb{C}^*)$, where $\mathcal{B} = \cup_{\nu\in\mathbb{N}}\mathcal{B}_\nu$, for $\mathcal{B}_\nu = z^{-1/\nu}\mathbb{C}[z^{-1/\nu}]$. In the non-formal case $K = \mathbb{C}(\{z\})$, the Galois group is given by the *Ramis wild fundamental group*. This has additional generators, which depend upon resummation of divergent series and are related to the Stokes phenomenon. We will not discuss this further here. We refer the interested reader to [**222**] and [**284**] for a more detailed treatment of both the formal and the non-formal case.

It is useful to remark, however, that there are suitable classes of differential systems for which the Stokes phenomena are not present, hence the differential Galois group is the same in the formal and in the non-formal setting (cf. e.g. Proposition 3.40 of [**284**]).

One recovers the differential Galois group of a single equation by considering the affine group scheme associated to the Tannakian subcategory of $\mathcal{D}_K$ generated by the differential module associated to the equation.

Similarly, one can restrict to particular Tannakian subcategories of $\mathcal{D}_K$ and identify them with categories of representations of affine group schemes. All these results can be seen as a particular case of the Riemann–Hilbert correspondence, which is broadly meant as an equivalence of categories between the analytic datum of certain equivalence classes of differential systems and a representation theoretic datum.

An important case is related to Hilbert's 21st problem, or the Riemann–Hilbert problem for regular–singular equations. For simplicity, we discuss here only the local case of equations over an infinitesimal disk $\Delta \subset \mathbb{C}$ around the origin. We refer the reader to [**105**] for a more general treatment.

EXAMPLE 1.91. *Let (K,δ) be the differential field $K = \mathbb{C}((z))$ with $\delta(f) = f'$. A differential equation* (1.342) *is regular-singular if there exists an invertible matrix T with coefficients in K such that $T^{-1}AT - T^{-1}\delta(T) = B/z$, where B is a matrix with coefficients in $\mathbb{C}[[z]]$. The Tannakian subcategory $\mathcal{D}_K^{\mathrm{rs}}$ of $\mathcal{D}_K$ generated by the regular-singular equations is equivalent to* $\mathrm{Rep}_{\bar{\mathbb{Z}}}$*, where $\bar{\mathbb{Z}}$ is the algebraic hull of $\mathbb{Z}$*

as in Example 1.86*. This corresponds to the fact that a regular singular equation is determined by its monodromy representation, where* $\mathbb{Z} = \pi_1(\Delta^*)$*.*

Deligne proved in [**105**] a much stronger form of the regular-singular Riemann–Hilbert correspondence, where instead of the simple local case of the infinitesimal disk one considers the general global case of a complement of divisors in an algebraic variety. In this case it is still true that the category of regular-singular differential systems is equivalent to the category of finite-dimensional linear representations of the fundamental group. This type of global regular-singular Riemann–Hilbert correspondence admits further generalizations, in the form of an equivalence of derived categories between regular holonomic $\mathcal{D}$-modules and perverse sheaves. We refer the reader to a short survey of this more general viewpoint given in §8 of [**139**].

Here we are interested specifically in identifying the Riemann–Hilbert correspondence for the equivalence classes of flat equisingular connections of §7.2. This case differs from the usual setting recalled here in the following ways.

- Our base space is not just the disk Δ but the $\mathbb{C}^*$-fibration B over Δ, so we leave the category $\mathcal{D}_K$ of differential systems of Definition 1.90.
- The equivalence relation on connections that we consider is through gauge transformations regular at $z = 0$.
- The equisingular connections are not regular-singular, hence we need to work in the setting of the "irregular" Riemann–Hilbert correspondence, where one allows for an arbitrary degree of irregularity, as in the case of $\mathcal{D}_K$ above.
- The Galois group will be the same in both the formal and the non-formal setting.

Recall that we already know from the proof of Theorem 1.67 that flat equisingular connections have trivial monodromy, so that the representation theoretic datum that classifies them is certainly not a monodromy representation. We see that the Galois group has a structure similar to the Ramis exponential torus.

We discuss all this in detail in the next section.

7.6. Universal Hopf algebra and the Riemann–Hilbert correspondence.

The main result of this section is the Riemann–Hilbert correspondence underlying the theory of perturbative renormalization. This will identify a universal affine group scheme $\mathbb{U}^*$ that governs the structure of the divergences of all renormalizable theories and which has the type of properties that Cartier envisioned for his "cosmic Galois group".

We begin by constructing a category of *flat equisingular vector bundles*. We will then see how it relates to the flat equisingular connections for a particular choice of the pro-unipotent affine group scheme G.

We first introduce the notion of a W-connection on a filtered vector bundle. As above we let $B^0 \subset B$ be the complement of the fiber over $z = 0 \in \Delta$ for the space B of (1.298).

DEFINITION 1.92. *Let (E, W) be a filtered vector bundle over B, with an increasing filtration $W^{-n-1}(E) \subset W^{-n}(E)$ and a given trivialization of the associated graded*

$$\mathrm{Gr}_n^W(E) = W^{-n}(E)/W^{-n-1}(E).$$

A W-connection on E is a connection ∇ on the vector bundle $E^0 = E|_{B^0}$, with the following properties :

(1) *The connection ∇ is compatible with the filtration, i.e. it restricts to all the $W^{-n}(E^0)$.*
(2) *The connection ∇ induces the trivial connection on the associated graded $\mathrm{Gr}^W(E)$.*

Two W-connections ∇_i on E^0 are W-equivalent iff there exists an automorphism T of E, preserving the filtration, that induces the identity on $\mathrm{Gr}^W(E)$ and conjugates the connections, $T \circ \nabla_1 = \nabla_2 \circ T$.

We reformulate conditions (1.300) and (1.301) of equisingularity in this context. Let V be a finite-dimensional $\mathbb{Z}$-graded vector space $V = \oplus_{n\in\mathbb{Z}} V_n$, and consider the trivial vector bundle $E = B \times V$. We view it as a $\mathbb{G}_m$-equivariant filtered vector bundle (E, W), with W the weight filtration

$$W^{-n}(V) = \oplus_{m \geq n} V_m, \tag{1.344}$$

and the $\mathbb{G}_m$-action induced by the grading of V.

DEFINITION 1.93. *A flat W-connection ∇ on E is equisingular if it is $\mathbb{G}_m$-invariant and for any fundamental system of solutions of $\nabla\eta = 0$ the associated isomorphism between restrictions of E to sections $\sigma : \Delta \to B$ with $\sigma(0) = y_0$ is regular.*

We construct a category $\mathcal{E}$ of flat equisingular vector bundles as follows.

DEFINITION 1.94. *The objects* $\mathrm{Obj}(\mathcal{E})$ *are given by data $\Theta = [V, \nabla]$. Here V is a finite-dimensional $\mathbb{Z}$-graded vector space and ∇ is an equisingular W-connection on the filtered bundle $E^0 = B^0 \times V$. The brackets $[V, \nabla]$ mean taking the W-equivalence class of the connection ∇. The data (E, ∇) are a flat equisingular vector bundle.*

The morphisms $\mathrm{Hom}_{\mathcal{E}}(\Theta, \Theta')$ *are linear maps $T : V \to V'$ that are compatible with the grading and satisfy the following condition. Consider on the bundle $(E' \oplus E)^*$ the W-connection*

$$\nabla_1 := \begin{pmatrix} \nabla' & 0 \\ 0 & \nabla \end{pmatrix}$$

and the W-connection ∇_2, which is the conjugate of ∇_1 by the unipotent matrix

$$\begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}. \tag{1.345}$$

Then the requirement for a morphism $T \in \mathrm{Hom}_{\mathcal{E}}(\Theta, \Theta')$ is that these two W-connections are W-equivalent on B,

$$\nabla_2 = \begin{pmatrix} \nabla' & T\nabla - \nabla' T \\ 0 & \nabla \end{pmatrix} \sim \nabla_1 = \begin{pmatrix} \nabla' & 0 \\ 0 & \nabla \end{pmatrix}. \tag{1.346}$$

It is worth explaining why it is necessary to use the direct sum $E' \oplus E$ and the matrices in condition (1.346) for the definition of morphisms. This is related to the fact that we are dealing with filtered spaces. The problem is that the category of filtered vector spaces, with morphisms that are linear maps respecting the filtration, is not an abelian category. In fact, suppose given a linear map $f : V \to V'$ of filtered vector spaces, such that $f(W^{-n}(V)) \subset W^{-n}(V')$. In general, there is no relation between the restriction to the image of f of the filtration of V' and the filtration on the quotient of V by $\mathrm{Ker}(f)$, while in an abelian category one would expect an isomorphism between $V/\mathrm{Ker}(f)$ and $\mathrm{Im}(f)$. Thus, we have to refine our class of morphisms to one for which this property is satisfied.

We first explain the relation of Definition 1.94 to the equivalence classes of flat equisingular G-connections considered in Sections 7.2 and 7.3.

LEMMA 1.95. *Let G be the affine group scheme dual to a graded connected Hopf algebra $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$, with finite-dimensional $\mathcal{H}_n$, as in §7.2. Let ϖ be a flat equisingular connection as in Definition 1.65. Let $\xi : G \to \mathrm{GL}(V)$ be a finite-dimensional linear graded representation of G. Then the data (ϖ, ξ) determine an element $\Theta \in \mathrm{Obj}(\mathcal{E})$. Equivalent ϖ determine the same Θ.*

PROOF. Notice that a graded representation of the pro-unipotent G in V can be described equivalently as a graded representation of the Lie algebra $\mathfrak{g} = \mathrm{Lie}\, G$ in V. Since the Lie algebra $\mathfrak{g}$ is positively graded, both representations are compatible with the weight filtration $W^{-n}(V) = \oplus_{m\geq n} V_m$. The induced representation on the associated graded $\mathrm{Gr}^W_n(V) = W^{-n}(V)/W^{-n-1}(V)$ is the identity. We let $\xi^* : G^* \to \mathrm{GL}(V)$ be the natural extension of ξ to $G^* = G \rtimes \mathbb{G}_m$ which when restricted to $\mathbb{G}_m$ is the grading of V. Consider the vector bundle E on B associated to the trivial principal G^*-bundle $\tilde{P}$ by the representation ξ^*. The connection ϖ is a flat equisingular G-connection in the sense of Definition 1.65 and by Lemma 1.73 it gives a flat equisingular connection on the G^*-bundle $\tilde{P}^0$ of (1.334). It induces a connection ∇ on $E^0 = E|_{B^0}$, which using Lemma 1.73 is a flat equisingular W-connection in the sense of Definition 1.93. If ϖ and ϖ' are flat equisingular G-connection on P^0 that are equivalent as in Definition 1.66, then the corresponding W-connections ∇ and ∇' are W-equivalent. ☐

Thus, the category $\mathcal{E}$ of flat equisingular vector bundles of Definition 1.94 provides a universal setting where equivalence classes of flat equisingular G-connections for varying G can be analyzed simultaneously. We will see later how we recover the classes of flat equisingular connections for a particular G from the general case.

We first analyze the main properties of the category $\mathcal{E}$. We begin by introducing the affine group scheme $\mathbb{U}^*$ and the universal singular frame $\gamma_{\mathbb{U}}$.

DEFINITION 1.96. *Let $\mathcal{L}_{\mathbb{U}} = \mathcal{F}(1, 2, 3, \cdots)_\bullet$ denote the free graded Lie algebra generated by elements e_{-n} of degree n, for each $n > 0$. Consider the Hopf algebra*

$$\mathcal{H}_{\mathbb{U}} := U(\mathcal{F}(1, 2, 3, \cdots)_\bullet)^\vee, \tag{1.347}$$

which is the graded dual of the universal enveloping algebra of $\mathcal{L}_{\mathbb{U}}$. We denote by $\mathbb{U}$ the affine group scheme associated to the commutative Hopf algebra $\mathcal{H}_{\mathbb{U}}$, and by $\mathbb{U}^*$ the semidirect product $\mathbb{U}^* = \mathbb{U} \rtimes \mathbb{G}_m$, with the action given by the grading.

As an algebra $\mathcal{H}_{\mathbb{U}}$ is isomorphic to the linear space of noncommutative polynomials in variables f_n, $n \in \mathbb{N}$, with the shuffle product.

The sum

$$e = \sum_{1}^{\infty} e_{-n}\,, \tag{1.348}$$

defines an element of the Lie algebra $\mathcal{L}_{\mathbb{U}}$ of $\mathbb{U}$. Since $\mathbb{U}$ is by construction a pro-unipotent affine group scheme we can lift e to a morphism

$$\mathbf{rg}\,:\,\mathbb{G}_a \to \mathbb{U}\,, \tag{1.349}$$

of affine group schemes from the additive group $\mathbb{G}_a$ to $\mathbb{U}$. This morphism will play an important role from the point of view of renormalization. In fact, we will see in Theorem 1.106 that it gives a universal and canonical lift of the renormalization group.

Definition 1.97. *The universal singular frame is the loop in the pro-unipotent Lie group* $\mathbb{U}(\mathbb{C})$ *given by the formula*

$$\gamma_{\mathbb{U}}(z,v) = \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(e)\,\frac{du}{u}}\,. \tag{1.350}$$

We can compute explicitly the coefficients of the universal singular frame as follows.

Proposition 1.98. *The universal singular frame is given by*

$$\gamma_{\mathbb{U}}(-z,v) = \sum_{n\geq 0, k_j>0} \frac{e_{-k_1}e_{-k_2}\cdots e_{-k_n}}{k_1\,(k_1+k_2)\cdots(k_1+k_2+\cdots+k_n)}\, v^{\sum k_j}\, z^{-n} \tag{1.351}$$

Proof. Using (1.348) and (1.262), we obtain that the coefficient of $e_{-k_1}e_{-k_2}\cdots e_{-k_n}$ is given by the expression

$$v^{\sum k_j}\, z^{-n} \int_{0\leq s_1\leq\cdots\leq s_n\leq 1} s_1^{k_1-1}\cdots\, s_n^{k_n-1}\, ds_1\cdots ds_n\,.$$

This yields the desired result. □

Remark 1.99. *Notice that the coefficients of* $\gamma_{\mathbb{U}}(z,v)$ *are all rational numbers. Moreover, the coefficients of* (1.351) *are the same coefficients that appear in the local index formula of* [**92**].

The main result of this section, which gives the Riemann–Hilbert correspondence for the category of flat equisingular bundles, is the following theorem.

Theorem 1.100. *Let* $\mathcal{E}$ *be the category of equisingular flat bundles of Definition 1.94. Let* $\omega : \mathcal{E} \to \mathrm{Vect}_{\mathbb{C}}$ *be the functor defined by* $\omega(\Theta) = V$*, for* $\Theta = [V,\nabla]$*. Then* $\mathcal{E}$ *is a neutral Tannakian category with fiber functor* ω *and is equivalent to the category of representations* $\mathrm{Rep}_{\mathbb{U}^*}$ *of the affine group scheme* $\mathbb{U}^*$ *of Definition* 1.96.

Proof. We prove this result in several steps. Let V be a finite-dimensional graded vector space over a field k. (Here we work with $k = \mathbb{C}$, but see Corollary 1.105 below.) Let $G_V(k)$ denote the group of linear transformations $S \subset \mathrm{End}(V)$ compatible with the weight filtration,

$$S\,W_{-n}(V) \subset W_{-n}(V), \tag{1.352}$$

and inducing the identity on the associated graded

$$S|_{\mathrm{Gr}_n^W} = 1. \tag{1.353}$$

The unipotent algebraic group G_V is (non-canonically) isomorphic to the unipotent group of upper triangular matrices. Its Lie algebra is then identified with strictly upper triangular matrices. We then have a direct translation between W-connections and G-valued connections.

PROPOSITION 1.101. *Let (E, ∇) be a flat equisingular vector bundle, with $E = B \times V$. Then the following holds.*

(1) *The connection ∇ defines a flat equisingular G_V-valued connection ϖ, for G_V as above.*
(2) *All flat equisingular G_V-valued connections are obtained in this way.*
(3) *This bijection preserves equivalence, namely equivalent flat equisingular G_V-valued connections correspond to the same object $\Theta = [V, \nabla]$ of $\mathcal{E}$.*

PROOF. Since W-connections are compatible with the filtration and trivial on the associated graded, they are obtained by adding a $\mathrm{Lie}\, G_V$-valued 1-form ϖ to the trivial connection. Similarly, W-equivalence is given by the equivalence of Definition 1.66. ☐

LEMMA 1.102. (1) *Let $\Theta = [V, \nabla]$ be an object of $\mathcal{E}$. There exists a unique representation $\rho = \rho_\Theta$ of $\mathbb{U}^*$ in V such that the restriction to $\mathbb{G}_m \subset \mathbb{U}^*$ is the grading and*

$$D\rho(\gamma_{\mathbb{U}}) \simeq \nabla, \tag{1.354}$$

where $\gamma_{\mathbb{U}}$ is the universal singular frame.

(2) *Given a representation $\rho : \mathbb{U}^* \to \mathrm{GL}(V)$, there exists a connection ∇ on $E^0 = B^0 \times V$, unique up to equivalence, such that $[V, \nabla]$ is an object in $\mathcal{E}$ and ∇ satisfies* (1.354).

PROOF. Let G_V be defined as above. By Proposition 1.101, we know that ∇ defines a flat equisingular G_V-valued connection ϖ. Then, using (1.303) of Theorem 1.67, we know that we have an equivalence

$$\varpi \sim D\left(\mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\beta)\,\frac{du}{u}}\right),$$

for a unique element $\beta \in \mathrm{Lie}\, G_V$. We can decompose the element β into homogeneous components for the action of the grading, $\beta = \sum_n \beta_n$ with $Y(\beta_n) = n\,\beta_n$. Thus, the element β (and the grading) uniquely determine a representation ρ of $\mathbb{U}^*$ in G_V, where $\rho(e_{-n}) = \beta_n$. This representation satisfies (1.354) by construction. Conversely, given a representation ρ of $\mathbb{U}^*$, we consider the grading associated to the restriction to $\mathbb{G}_m$ and let

$$\gamma(z, v) = \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\rho(e))\,\frac{du}{u}}.$$

The flat equisingular connection $D\gamma$ determines a W-connection ∇ on the vector bundle E^0 with the desired properties. ☐

We will see in Theorem 1.106 below that the construction of representations of $\mathbb{U}$ in G_V used in the proof of Lemma 1.102 above holds in general and provides the way to recover classes of flat equisingular connections for a given pro-unipotent affine group scheme G from the data of the category $\mathcal{E}$. We now continue with the proof of Theorem 1.100.

LEMMA 1.103. *Let $\Theta = [V, \nabla]$ be an object in $\mathcal{E}$. Then the following holds.*

(1) *For any $S \in \mathrm{Aut}(V)$ compatible with the grading, $S \nabla S^{-1}$ is an equisingular connection.*
(2) *The representation ρ_Θ of Lemma 1.102 satisfies*

$$\rho_{[v, S \nabla S^{-1}]} = S\, \rho_{[V,\nabla]}\, S^{-1}.$$

(3) *The equisingular connections ∇ and $S \nabla S^{-1}$ are equivalent if and only if $[\rho_{(E,\nabla)}, S] = 0$.*

PROOF. (1) The equisingularity condition is satisfied. In fact, the $\mathbb{G}_m$-invariance follows from the compatibility with the grading and the restriction to a section $\sigma : \Delta \to B$ satisfies

$$\sigma^*(S \nabla S^{-1}) = S\, \sigma^*(\nabla)\, S^{-1},$$

so that the second condition for equisingularity is also satisfied.
(2) The second statement follows from Lemma 1.102 and the compatibility of S with the grading. In fact, we have, for an element $\beta \in \mathrm{Lie}\, G_V$,

$$S\, \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\beta)\,\frac{du}{u}}\, S^{-1} = \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(S\,\beta\, S^{-1})\,\frac{du}{u}}.$$

(3) The third statement follows immediately from the second, since equivalence corresponds to having the same $\beta \in \mathrm{Lie}\, G_V$, by Theorem 1.67. $\square$

PROPOSITION 1.104. *Let $\Theta = [V, \nabla]$ and $\Theta' = [V', \nabla']$ be objects of $\mathcal{E}$. Let $T : V \to V'$ be a linear map compatible with the grading. Then the following two conditions are equivalent:*

(1) *$T \in \mathrm{Hom}_{\mathcal{E}}(\Theta, \Theta')$;*
(2) *$T \rho_\Theta = \rho_{\Theta'}\, T$.*

PROOF. Let $S \in \mathrm{Aut}(E' \oplus E)$ be the unipotent automorphism of (1.345). By construction, S is an automorphism of $E' \oplus E$, compatible with the grading. By (3) of the previous Lemma, we have

$$S \begin{pmatrix} \nabla' & 0 \\ 0 & \nabla \end{pmatrix} S^{-1} \sim \begin{pmatrix} \nabla' & 0 \\ 0 & \nabla \end{pmatrix}$$

if and only if

$$\begin{pmatrix} \beta' & 0 \\ 0 & \beta \end{pmatrix} S = S \begin{pmatrix} \beta' & 0 \\ 0 & \beta \end{pmatrix}.$$

This holds if and only if $\beta' T = T \beta$. $\square$

To complete the proof of Theorem 1.100, we check that the functor $\mathrm{Rep}_{\mathbb{U}^*} \to \mathcal{E}$ given by $\rho \mapsto D\rho(\gamma_U)$ is compatible with the tensor structure. In $\mathcal{E}$ this is given by

$$(V, \nabla) \otimes (V', \nabla') = (V \otimes V', \nabla \otimes 1 + 1 \otimes \nabla').$$

This is compatible with the W-equivalence and with the condition of equisingularity of the connections. The compatibility of $\rho \mapsto D\rho(\gamma_U)$ with tensor products then follows from the formula

$$\mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\beta\otimes 1 + 1\otimes\beta')\,\frac{du}{u}} = \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\beta)\,\frac{du}{u}} \otimes \mathrm{T}e^{-\frac{1}{z}\int_0^v u^Y(\beta')\,\frac{du}{u}}.$$

On morphisms, it is sufficient to check the compatibility on $1 \otimes T$ and $T \otimes 1$.

We have shown that the tensor category $\mathcal{E}$ is equivalent to the category $\mathrm{Rep}_{\mathbb{U}^*}$ of finite-dimensional representations of $\mathbb{U}^*$. The fact that $\mathbb{U}^* = \underline{\mathrm{Aut}}^{\otimes}(\omega)$ then follows, since the fiber functor ω becomes through this equivalence the forgetful functor that assigns to a representation the underlying vector space.

$\square$

COROLLARY 1.105. *The category $\mathcal{E}$ can also be defined over the field $k = \mathbb{Q}$ and the equivalence of rigid k-linear tensor categories proved in Theorem* 1.100 *still holds for $k = \mathbb{Q}$.*

PROOF. In the case where $k = \mathbb{Q}$, we work in the formal setting with the differential field $K = \mathbb{Q}((z))$. We still consider the same geometric setting as before, where the infinitesimal disk $\Delta = \mathrm{Spec}(K)$. All the arguments still go through. In fact, notice that (1.273) gives a rational expression for the operator D. This, together with the fact that the coefficients of the universal singular frame in (1.351) are rational, implies that we can work with a rational ∇. $\square$

In particular, Corollary (1.105) also shows that for the category of flat equisingular bundles the formal and the nonformal theory give the same Galois group $\mathbb{U}^*$. This reflects the fact that, due to the pro-unipotent nature of the affine group scheme, our arguments usually depend upon only finitely many terms in an infinite sum, cf. Remark 1.59.

In the rational case, we can define, for each $n \in \mathbb{Z}$, an object $\mathbb{Q}(n)$ of the category $\mathcal{E}$ of equisingular flat vector bundles where V is given by a one-dimensional $\mathbb{Q}$-vector space placed in degree n, and ∇ is the trivial connection on the associated vector bundle E over B. Then the fiber functor takes the form $\omega = \oplus\, \omega_n$, with

$$\omega_n(\Theta) = \mathrm{Hom}(\mathbb{Q}(n), \mathrm{Gr}^W_{-n}(\Theta))\,. \tag{1.355}$$

We now return to discuss the physical significance of the result of Theorem 1.100 and the role of the affine group scheme $\mathbb{U}^*$ in perturbative renormalization. The following result shows how $\mathbb{U}^*$ is related to the pro-unipotent affine group scheme G. In particular, this holds in the case where $G = \mathrm{Difg}(\mathcal{T})$ is the group of diffeographisms of a renormalizable theory.

THEOREM 1.106. *Let G be a pro-unipotent affine group dual to a graded connected commutative Hopf algebra $\mathcal{H} = \oplus_{n\geq 0}\mathcal{H}_n$, with finite-dimensional $\mathcal{H}_n$. Then the following properties hold.*

(1) *There exists a canonical bijection between equivalence classes of flat equisingular connections on $\tilde{P}^0$ and graded representations $\mathbb{U} \to G$, or equivalently representations*

$$\rho : \mathbb{U}^* \to G^* = G \rtimes \mathbb{G}_m, \tag{1.356}$$

which are the identity on $\mathbb{G}_m$.

(2) *The universal singular frame $\gamma_{\mathbb{U}}$ provides universal counterterms. Namely, given a loop $\gamma_\mu \in L(G(\mathbb{C}), \mu)$, the universal singular frame maps to $\gamma_-(z)$ under the representation ρ of* (1.356)*.*

(3) *The renormalization group F_t in $G(\mathbb{C})$ described in Proposition 1.47 is obtained as the composite $\rho \circ \mathbf{rg}$, with ρ as in* (1.356) *and $\mathbf{rg} : \mathbb{G}_a \to \mathbb{U}$ as in* (1.349)*.*

PROOF. (1) Theorem 1.67 shows that the equivalence classes of flat equisingular connections are parameterized by the elements $\beta \in \mathrm{Lie}\, G$. We then proceed as we did in the proof of Lemma 1.102. The positivity and integrality of the grading make it possible to write β as an infinite formal sum

$$\beta = \sum_1^\infty \beta_n\,, \tag{1.357}$$

where, for each n, β_n is homogeneous of degree n for the grading, i.e. $Y(\beta_n) = n\beta_n$. Assigning β and the action of the grading on it is the same as giving a collection of homogeneous elements β_n that fulfill no restriction besides $Y(\beta_n) = n\beta_n$. In particular, there is no condition on their Lie brackets. Thus, assigning such data is the same as giving a graded homomorphism from the affine group scheme $\mathbb{U}$ to G, i.e. a ρ as in (1.356).

(2) This follows from Theorem 1.58, since we have

$$\rho(\gamma_{\mathbb{U}})(z,v) = \mathrm{T}e^{\frac{-1}{z}\int_0^v u^{Y}(\beta)\frac{du}{u}}\;.$$

(3) By Corollary 1.49 we know that β is the infinitesimal generator of F_t □

COROLLARY 1.107. *The affine group scheme* $\mathbb{U}$ *acts on the coupling constants of physical theories, through the representation* (1.356) *to the group of diffeographisms* $\mathrm{Difg}(\mathcal{T})$ *and then applying Proposition* 1.42 *to obtain*

$$\mathbb{U} \to \mathrm{Difg}(\mathcal{T}) \to \mathrm{Diff}. \tag{1.358}$$

Thus, the affine group scheme $\mathbb{U}$ has the right properties that Cartier expected for his "cosmic Galois group". In §8 we will also discuss the arithmetic nature of $\mathbb{U}$ and its relation to motivic Galois groups. Note that, while the morphism $\mathrm{Difg}(\mathcal{T}) \to \mathrm{Diff}$ is defined over $\mathbb{Q}$ the morphism $\mathbb{U} \to \mathrm{Difg}(\mathcal{T})$ in (1.358) is only defined over $\mathbb{C}$ since it is associated to an equisingular connection with complex coefficients. In the dual map of Hopf algebras, the image of the generators Γ will have coefficients with respect to the generators of $\mathcal{H}_{\mathbb{U}}$ that are transcendental numbers obtained from the residues of the graphs.

The map (1.356) can be seen as a map of Galois groups, from the Galois group of the Tannakian category $\mathcal{E}$ to that of the Tannakian subcategory generated by the flat equisingular vector bundles that come from flat equisingular G-connections on $\tilde{P}^0$. In general, for a given quantum field theory $\mathcal{T}$ the subcategory $\mathcal{E}_{\mathcal{T}}$ of $\mathcal{E}$ given by flat equisingular vector bundles coming from $\mathcal{T}$ will differ from $\mathcal{E}$. The corresponding affine group scheme plays the same role as the Galois group of a given differential equation does with respect to the universal group. Thus it is natural to define the Galois group of $\mathcal{T}$ as follows, with

$$\mathrm{Gal}(\mathcal{T})^* = \mathrm{Gal}(\mathcal{T}) \rtimes \mathbb{G}_m. \tag{1.359}$$

DEFINITION 1.108. *Let* $\mathcal{T}$ *be a given renormalizable quantum field theory. We define the Galois group* $\mathrm{Gal}\,(\mathcal{T})^*$ *as the affine group scheme associated to the subcategory* $\mathcal{E}_{\mathcal{T}}$ *of flat equisingular vector bundles coming from* $\mathcal{T}$*. We let* $\mathrm{Gal}\,(\mathcal{T})$ *be the graded affine group scheme kernel of the canonical morphism* $\mathrm{Gal}\,(\mathcal{T})^* \to \mathbb{G}_m$.

Since we are dealing with a subcategory of $\mathcal{E}$, one has a natural surjection

$$\mathbb{U} \to \mathrm{Gal}\,(\mathcal{T}) \tag{1.360}$$

and the group homomorphism $\rho : \mathbb{U} \to G$ of (1.356) with $G = \mathrm{Difg}(\mathcal{T})$ factors through $\mathrm{Gal}\,(\mathcal{T})$ by (1.360). When there are no divergences in the theory one has $\mathrm{Gal}\,(\mathcal{T}) = \{1\}$ and this can happen even though $\mathrm{Difg}(\mathcal{T})$ is quite large. It is of course desirable to find in our framework an analogue of Theorem 1.89 on the resolution of equations.

PROPOSITION 1.109. *Let $\mathcal{T}$ be a given renormalizable quantum field theory.*
1) The theory $\mathcal{T}$ is finite iff its Galois group is trivial, that is $\mathrm{Gal}\,(\mathcal{T}) = \{1\}$.
2) If the theory $\mathcal{T}$ is super-renormalizable, its Galois group $\mathrm{Gal}\,(\mathcal{T})$ *is finite-dimensional.*

PROOF. 1) If the theory is finite then there are no divergences and $\beta = 0$ so that the image of $\mathbb{U}$ in $G = \mathrm{Difg}(\mathcal{T})$ is trivial. The converse also holds, since the divergences are determined by the residues and hence by β.
2) By definition (cf. [**62**] §5.7.3) a theory is super-renormalizable iff only a finite number of graphs need overall counterterms. The prototypical example is $\mathcal{T} = \phi_4^3$. It then follows that only a finite number of homogeneous components β_k are non-zero so that the Lie algebra of $\mathrm{Gal}\,(\mathcal{T})$ is finite-dimensional. □

In particular, $\mathrm{Lie}\,\mathrm{Gal}(\mathcal{T})$ is the Lie subalgebra of $\mathrm{Lie}\,\mathrm{Difg}(\mathcal{T})$ generated by the β_n.

8. Motives in a nutshell

In this section we give a brief overview of Grothendieck's theory of motives. The theory is vast and one can easily get into very hard technical aspects, but we will only give an impressionistic sketch, and focus on those aspects that are of direct relevance to the interaction with noncommutative geometry and quantum physics. For the reader who might be interested in a more detailed treatment of the subject, we recommend the recent book by Yves André [**2**], and the two AMS volumes of the 1991 Seattle conference (cf. [**173**]).

8.1. Algebraic varieties and motives.

We describe in this section those aspects of the theory of algebraic varieties that contributed to the formulation of Grothendieck's original idea of a theory of motives.

8.1.1. *Cohomology theories.* With the development of étale cohomology of algebraic varieties, one finds that there are several different viable cohomology theories one can associate to an algebraic variety. Moreover, these theories are related by specific comparison isomorphisms.
If X is a smooth projective algebraic variety over a field k, then for a given separable closure k^{sep} and for $\bar{X}$ the variety over k^{sep} obtained by extension of scalars, one can define the ℓ-adic étale cohomology

$$H^i_{\mathrm{et}}(\bar{X}, \mathbb{Q}_\ell) \tag{1.361}$$

whenever ℓ is prime to the characteristic of k. These are finite-dimensional $\mathbb{Q}_\ell$-vector spaces, which satisfy all the desirable properties of a cohomology theory: Poincaré duality, Künneth formula, Lefschetz formula. Moreover, the spaces $H^1_{\mathrm{et}}(\bar{X}, \mathbb{Q}_\ell)$ carry a representation of the absolute Galois group $\mathrm{Gal}(\bar{k}/k)$.

In the case where the field k is of characteristic zero, one can also define de Rham cohomology groups

$$H^i_{\mathrm{dR}}(X, k) = \mathbb{H}^i(\Omega^*_{X/k}), \tag{1.362}$$

which are finite-dimensional filtered k-vector spaces. These are constructed by taking the hypercohomology of the algebraic de Rham complex $\Omega^*_{X/k}$.

If the field k can be embedded in $\mathbb{C}$, then the choice of such an embedding $\sigma : k \hookrightarrow \mathbb{C}$ determines another cohomology theory, the Betti cohomology

$$H^i_{\mathrm{B}}(X, \mathbb{Q}) := H^i(X(\mathbb{C}), \mathbb{Q}) \tag{1.363}$$

is the singular cohomology of $X(\mathbb{C})$. These are finite-dimensional $\mathbb{Q}$-vector spaces. They have the property that the $\mathbb{C}$-vector space $H^i_{\mathrm{B}}(X, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{C}$ is bigraded (it carries a Hodge structure)

$$H^i_{\mathrm{B}}(X, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{C} = \oplus_{p+q=i} H^{p,q}.$$

There are interesting isomorphisms relating these cohomology theories. The period isomorphism relates Betti and de Rham cohomology

$$H^i_{\mathrm{B}}(X, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{C} \simeq H^i_{\mathrm{dR}}(X, k) \otimes_k \mathbb{C}. \tag{1.364}$$

The matrix realizing the period isomorphism is the *period matrix*.

The comparison isomorphisms relate étale and Betti cohomology by

$$H^i_{\mathrm{B}}(X, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \simeq H^i_{\mathrm{et}}(\bar{X}, \mathbb{Q}_\ell). \tag{1.365}$$

In positive characteristic and for a (perfect) field k, one can also consider crystalline cohomology. This gives a functor from varieties over k to graded $W(k)$-algebras that are finitely generated as $W(k)$-modules. Here $W(k)$ is the ring of Witt vectors of k. Again there is a comparison isomorphism, which states that if X/k lifts to a smooth $\mathcal{X}/W(k)$ then crystalline and de Rham cohomology agree

$$H^i_{\mathrm{crys}}(X/k) \simeq H^i_{\mathrm{dR}}(\mathcal{X}/W(k)) = \mathbb{H}^i(\Omega^*_{\mathcal{X}/W(k)}). \tag{1.366}$$

One can see a very simple example of these different cohomologies by looking at $H^2(\mathbb{P}^1)$. This particular cohomology will play a role below, when we talk about Tate motives. One has the following ([**2**], p.29)

- $H^2_{\mathrm{et}}(\mathbb{P}^1, \mathbb{Q}_\ell) = \mathrm{Hom}_{\mathbb{Z}_\ell}(\varprojlim_n \mu_{\ell^n}, \mathbb{Q}_\ell)$ (étale): μ_m is the group of m-th roots of unity.
- $H^2_{\mathrm{dR}}(\mathbb{P}^1, k) = k$ (de Rham) with the filtration $F^\bullet$ with $F^{\leq 0} = 0$ and $F^{>0} = k$.
- $H^2_{\mathrm{B}}(\mathbb{P}^1, \mathbb{Q}) = \frac{1}{2\pi i}\mathbb{Q}$ (Betti).
- $H^2_{\mathrm{crys}}(\mathbb{P}^1) = W(k)[1/p]$ (crystalline).

One can abstract from these cases (de Rham, étale, Betti, crystalline) the desirable properties of a "good" cohomology theory of algebraic varieties. This leads to the notion of Weil cohomology.

A Weil cohomology H^* is a contravariant functor from the category $\mathcal{V}(k)$ of smooth projective varieties over a field k (or from the category of smooth projective schemes) to the category GrVect(K) of graded vector spaces over a field K of characteristic zero. The required properties for a Weil cohomology are summarized as follows.

- The vector spaces $H^i(X)$ are finite-dimensional in each degree i, with $H^i(X) = 0$ unless $0 \leq i \leq 2\dim X$.
- Poincaré duality holds, namely for each $X \in \mathrm{Obj}(\mathcal{V}(k))$ there is an isomorphism $H^{2\dim X}(X) \simeq K$ and a nondegenerate bilinear pairing $(x, y) \mapsto \langle x, y\rangle$,

$$H^i(X) \times H^{2\dim X - i}(X) \stackrel{\simeq}{\to} K. \tag{1.367}$$

- The Künneth formula holds: the projections $\pi_i : X_1 \times X_2 \to X_i$, for $i = 1, 2$, induce an isomorphism

$$H^*(X \times_k Y) \simeq H^*(X) \otimes_K H^*(Y). \tag{1.368}$$

- If $C^i(X)$ denotes the abelian group of algebraic cycles (linear combinations of irreducible closed subvarieties) in X of codimension i, there is a cycle map

$$\gamma^i_X : C^i(X) \to H^{2i}(X) \tag{1.369}$$

which satisfies functoriality properties under pullback (and pushforwards by Poincaré duality), has the multiplicative property

$$\gamma^{i+j}_{X\times Y}(Z \times W) = \gamma^i_X(Z) \otimes \gamma^j_Y(W)$$

and is normalized by the condition that γ_{pt} is the inclusion $\mathbb{Z} \subset K$.
- Weak and hard Lefschetz conditions are satisfied. Namely if W is a smooth hyperplane section of X and $\iota : W \to X$ is the inclusion then the induced map $\iota^* : H^i(X) \to H^i(W)$ is an isomorphism for $i \leq \dim X - 2$ and is injective for $i = \dim X - 1$. This is the weak Lefschetz condition, while hard Lefschetz states that there are isomorphisms

$$\mathcal{L}^{(\dim X - i)} : H^i(X) \stackrel{\simeq}{\to} H^{2\dim X - i}(X), \tag{1.370}$$

for all $i \leq \dim X$, where $\mathcal{L}$ is the Lefschetz operator

$$\mathcal{L} : H^i(X) \to H^{i+2}(X), \quad \mathcal{L}(x) = x \cup \gamma^1_X(W).$$

The classical cohomology theories mentioned above (Betti, de Rham, étale, crystalline) are Weil cohomologies. For example, consider the case of a field k of characteristic zero. Betti cohomology is a Weil cohomology because singular cohomology for varieties over $\mathbb{C}$ is a Weil cohomology. By results of Grothendieck, Artin, and finally Deligne's proof of the hard Lefschetz properties, it is known that étale cohomology is a Weil cohomology.

Notice that the existence of the cycle map (1.369) for a Weil cohomology implies that there is always a part of the cohomology (the image of the cycle map) that is realized by algebraic cycles. In particular, this shows that correspondences given by algebraic cycles in the product of two varieties determine morphisms of any Weil cohomology through the cycle map. This is the key to defining morphisms in the category of motives. In general, one has very little explicit knowledge of which part of a given cohomology theory lies in the image of the cycle map (the Hodge conjecture being an example of this type of question).

The natural question prompted by the existence of so many different cohomology theories of algebraic varieties and of comparison isomorphisms between them is to understand what type of information can be transferred from one cohomology theory to another. The fundamental idea of motives is the existence of a universal cohomology theory for algebraic varieties, with values not in vector spaces but in a suitable $\mathbb{Q}$-linear semi-simple monoidal category (see below), which maps to all the "reasonable" cohomology theories (i.e. to all Weil cohomologies). A cohomological invariant is motivic if it exists at this universal level and it manifests itself in the various cohomologies through suitable realization functors.

8.1.2. *Algebraic cycles.* The fundamental step in constructing such a universal cohomology theory is a "linearization" of the category of algebraic varieties. This is achieved by replacing the usual notion of morphisms of varieties by a larger class of morphisms given by correspondences realized by algebraic cycles in the product.
Let $\mathcal{V}(k)$ be the category of smooth projective algebraic varieties over a field k. It is not an additive category, but it can be embedded in an additive category with the same objects and with morphisms given by formal linear combinations (with coefficients in $\mathbb{Z}$ or in $\mathbb{Q}$) of algebraic cycles in the product $X \times Y$ of codimension equal to the dimension of X.
Thus, a correspondence Z is of the form $Z = \sum_i n_i Z_i$, where the Z_i are irreducible algebraic subvarieties of $X \times Y$ of the prescribed codimension. We denote the resulting abelian group (or $\mathbb{Q}$-vector space) by $C^0(X,Y)_{\mathbb{Z}}$ (or $C^0(X,Y)_{\mathbb{Q}}$).
Notice that, as in the case of a cohomology theory, we are constructing a contravariant functor from varieties to motives, so that the abelian group $C^0(X,Y)_{\mathbb{Z}}$ of algebraic cycles in $X \times Y$ corresponds to maps from Y to X, hence the condition of the codimension being equal to $\dim X$. A morphism of varieties $f : Y \to X$ determines a morphism in this generalized sense by taking the (transpose of the) graph $\Gamma(f)^t \subset X \times Y$.

It is in fact better to consider morphisms as defined by algebraic cycles modulo a certain equivalence relation. For the purposes of enumerative algebraic geometry, it suffices to consider cycles up to the *numerical* equivalence relation. This states that two cycles $Z_1, Z_2 \in C^0(X,Y)$ are equivalent, $Z_1 \sim_{\rm num} Z_2$, iff the intersection numbers

$$Z_1 \bullet Z = Z_2 \bullet Z \tag{1.371}$$

are the same, for any Z a subvariety of complementary dimension. The intersection number is defined by linearity: if $Z_i = \sum_j n_{ij} Z_{ij}$ then $Z_i \bullet Z = \sum_j n_{ij} Z_{ij} \bullet Z$.

The category of numerical effective motives $\mathcal{M}^{\rm eff}_{\rm num}(k)$ is the pseudo-abelian envelope (the minimal pseudo-abelian category spanned by an additive category, it also called the Karoubian envelope) of the category generated by the algebraic varieties with morphisms $C^0(X,Y)_{\mathbb{Q},{\rm num}} = C^0(X,Y)_{\mathbb{Q}}/\sim_{\rm num}$. More precisely, this means that the objects of $\mathcal{M}^{\rm eff}_{\rm num}(k)$ are pairs (X,p) of $X \in {\rm Obj}(\mathcal{V}(k))$ and p a projector, that is, an endomorphism

$$p \in C^0(X,X)_{\mathbb{Q},{\rm num}}$$

satisfying $p^2 = p$ (in the numerical equivalence relation). Morphisms between two such objects are of the form

$$\mathrm{Hom}_{\mathcal{M}^{\rm eff}_{\rm num}(k)}((X,p),(Y,q)) = q \circ C^0(X,Y)_{\mathbb{Q},{\rm num}} \circ p. \tag{1.372}$$

An important result of Jannsen [**174**] shows that the category $\mathcal{M}^{\rm eff}_{\rm num}(k)$ is a semi-simple abelian category. It is endowed with a monoidal structure induced by the Cartesian product of varieties.

Another notation for elements (X, p) in the category $\mathcal{M}^{\rm eff}_{\rm num}(k)$ is $p\,h(X)$. In particular, one often writes $h^i(X)$ for an element (X, p^i) in $\mathcal{M}^{\rm eff}_{\rm num}(k)$ that maps to $H^i(X)$ in the realization in a Weil cohomology (cf. § 8.2.4 below on the standard conjectures). The decomposition $H^*(X) = \oplus_i H^i(X)$ as a graded vector space corresponds to the fact that an algebraic variety has an additive decomposition in the category of motives of the form $h(X) = \oplus_i h^i(X)$. In this sense, motives can be thought of as "direct summands of algebraic varieties". One says that the motive $h^i(X)$ is pure of weight i.

One can consider other choices of the equivalence relation on algebraic cycles. One choice is *rational* equivalence. In this case one says that $Z_1 \sim_{\rm rat} Z_2$ if they can be transformed one into the other by a sequence of deformations parameterized by the projective line $\mathbb{P}^1$. The group of algebraic cycles modulo rational equivalence is the Chow group $A^i(X)$. One denotes the resulting vector space of correspondences by $C^0(X, Y)_{\mathbb{Q},\rm rat}$ and one obtains in this way a category of effective Chow motives $\mathcal{M}^{\rm eff}_{\rm rat}(k)$ as above, by replacing $\sim_{\rm num}$ with $\sim_{\rm rat}$. As we discuss below, these categories have different properties when one changes the choice of the equivalence relation.

Another important equivalence relation on algebraic cycles is the homological relation. This depends upon the choice of a Weil cohomology H^*. As we have seen, such a cohomology comes endowed with a cycle map $\gamma^i : A^i(X) \to H^{2i}(X)(i)$. We say that an algebraic cycle $Z \in C^0(X, Y)$ is homologically trivial if its image under the cycle map is trivial, $\gamma(Z) = 0$. One denotes the resulting equivalence relation by $\sim_{H-\rm hom}$, or $\sim_{\rm hom}$ when the choice of H^* is understood. One can then define a category of effective motives up to homological equivalence $\mathcal{M}^{\rm eff}_{\rm hom}(k)$ as above.

The rational equivalence relation is the finest among those mentioned, while the numerical is the coarsest. It is part of the *standard conjectures* (which we discuss briefly later in this section) that the numerical and the homological equivalence relations are expected to coincide. In particular the homological relation is independent of the choice of the Weil cohomology.

We denote by $C^r(X, Y)_{\mathbb{Q}}$ the vector space of algebraic cycles of codimension $\dim X + r$. The composition law for correspondences given by algebraic cycles is defined using the intersection product. For $Z \in C^0(X_1, X_2)$ and $W \in C^0(X_2, X_3)$ one defines the composition as

$$W \circ Z = (\pi_{13})_* \left(\pi_{12}^* Z \bullet \pi_{23}^* W\right), \tag{1.373}$$

where $\pi_{ij} : X_1 \times X_2 \times X_3 \to X_i \times X_j$ are the projections. The law (1.373) induces an associative composition product

$$\circ : C^r(X_1, X_2)_{\mathbb{Q},\sim} \times C^s(X_2, X_3)_{\mathbb{Q},\sim} \to C^{r+s}(X_1, X_3)_{\mathbb{Q},\sim}$$

for any of the equivalence relations mentioned above.

8.1.3. *Zeta functions.* Another important input that contributed to the origin of motives is the question of counting points of varieties over finite fields, and generalizations in terms of L-functions of varieties.

Given a variety X defined over a finite field $\mathbb{F}_q$, with $q = p^\ell$ for a prime p, one wants to count points of X over the fields $\mathbb{F}_{q^n}$. It is natural to introduce a generating function for the numbers $N_n = \#X(\mathbb{F}_{q^n})$ in the form

$$\log Z_X(t) = \sum_n \frac{N_n}{n}\, t^n. \tag{1.374}$$

The numbers N_n have an interpretation as intersection numbers of the diagonal in $X \times X$ with the correspondence given by the graph of the n-th power Fr_X^n of the Frobenius endomorphism induced by the transformation $x \mapsto x^q$ on the algebra of functions on X. Thus, the question of computing such intersection numbers can be seen as a question naturally formulated in a category of motives up to numerical equivalence.

In the topological context, the question of counting the number of isolated fixed points of a map (intersections of the graph with the diagonal) can be treated efficiently using the Lefschetz fixed point formula, which turns it onto a cohomological formulation

$$\#\mathrm{Fix}(f) = \sum_i (-1)^i \mathrm{Tr}(f^* | H^i(X)). \tag{1.375}$$

where $f^*|H^i(X)$ is the action of the induced map on cohomology.

André Weil pioneered the use of this method in the case of the counting of points $X(\mathbb{F}_{q^n})$. In fact, if one has a suitable cohomology theory H^* for algebraic varieties over finite fields for which a version of the Lefschetz fixed point formula applies, then one can reformulate the counting of $N_n = \#X(\mathbb{F}_{q^n})$ in terms of a Lefschetz formula for the Frobenius endomorphism

$$N_n = \sum_i (-1)^i \mathrm{Tr}((Fr_X^n)^* | H^i(X)), \tag{1.376}$$

so that the zeta function $Z_X(t)$ can be expressed in the form

$$Z_X(t) = \prod_{i=0}^{2\dim X} \det(1 - t\, Fr_X^* | H^i(X))^{(-1)^{i+1}}. \tag{1.377}$$

A first glimpse of why the decomposition as a product over the odd or even cohomology groups $H^i(X)$ of the zeta function can be thought of as a motivic phenomenon comes from the very simple formula

$$\#\mathbb{P}^n(\mathbb{F}_q) = 1 + q + q^2 + \cdots + q^n,$$

which can be seen as a manifestation of the fact that the variety $\mathbb{P}^n$ has a decomposition in the category of motives of the form

$$h(\mathbb{P}^n) = 1 \oplus L \oplus L^2 \oplus \cdots \oplus L^n,$$

where 1, also denoted by $\mathbb{Q}(0)$, is the motive of a point, and L is the Lefschetz motive, which is also denoted by $\mathbb{Q}(-1)$. It is the motive $L = h^2(\mathbb{P}^1)$. Thus, one thinks in general of the product decomposition of (1.377) as a motivic decomposition

$$Z_X(t) = \prod_{i=0}^{2\dim X} L(h^i(X), t)^{(-1)^{i+1}},$$

into L-factors $L(h^i(X), t)$ associated to the motives $h^i(X)$.

This approach and this class of zeta functions was then extended to define L-functions for smooth projective varieties over number fields, which can be thought of as a far-reaching generalization of the simplest example, the Riemann zeta function

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}. \tag{1.378}$$

This can be thought of as the L-function of the zero-dimensional variety $\operatorname{Spec}\mathbb{Z}$. More generally, suppose that X is an algebraic variety defined over a number field $\mathbb{K}$. Then the associated Hasse–Weil L-function is given in the form

$$L(X,s) = \prod_i L(H^i(X),s)^{(-1)^{i+1}}. \tag{1.379}$$

Each L-function $L(H^i(X),s)$ in turn can be written as an Euler product over the finite places of $\mathbb{K}$,

$$L(H^i(X),s) = \prod_{\wp\in\operatorname{Spec}(\mathcal{O}_{\mathbb{K}})} L_\wp(H^i(X),s), \tag{1.380}$$

where the Euler factor $L_\wp(H^i(X),s)$ is associated to the reduction of X modulo the prime $\wp$ in the following way.

In general, for $\mathbb{K}$ a global field and $\mathbb{K}_v$ the local field given by the completion in the valuation associated to a non-Archimedean place v of $\mathbb{K}$, the Galois group $G_v = \operatorname{Gal}(\mathbb{K}_v^{\rm sep}/\mathbb{K})$, for $\mathbb{K}_v^{\rm sep}$ a separable closure, contains the inertia subgroup

$$I_v = \operatorname{Gal}(\mathbb{K}_v^{\rm sep}/\mathbb{K}_v^{\rm un}),$$

with $\mathbb{K}_v^{\rm un}$ the maximal unramified extension of $\mathbb{K}_v$ in $\mathbb{K}_v^{\rm sep}$. If the residue field k_v is a finite field, then the quotient G_v/I_v has a canonical generator given by the Frobenius. Given a vector space V with a representation of G_v, the Frobenius induces an automorphism of the inertia invariants V^{I_v}. The inverse of this automorphism is called the geometric Frobenius acting on V^{I_v}. In particular, for X_v a smooth projective variety over $\mathbb{K}_v$, one has an induced geometric Frobenius on the inertia invariants of the ℓ-adic étale cohomology of $\bar{X}_v = X_v \times_{\mathbb{K}_v} \mathbb{K}_v^{\rm sep}$, with ℓ prime to the characteristic of k_v.

The local Euler factor $L_\wp(H^i(X),s)$ is of the form

$$L_\wp(H^i(X),s) = \det(1 - Fr_\wp^*\,\mathfrak{n}(\wp)^{-s}\,|\,H^i(\bar{X}_\wp,\mathbb{Q}_\ell)^{I_\wp})^{-1}, \tag{1.381}$$

where $\mathfrak{n}(\wp) = \#k_\wp$ is the cardinality of the residue field at $\wp$ and $Fr_\wp^*$ is the induced geometric Frobenius. (See [**264**] for more details.)

An important advantage of passing from the combined L-function (1.379) to a single factor $L(H^i(X),s)$ is that one avoids the problem of infinitely many poles arising from the zeros of the factors in the denominator. Moreover, at least conjecturally, the $L(H^i(X),s)$ are expected to have a nice functional equation, when one completes the product (1.380) with a suitable contribution from the Archimedean places of $\mathbb{K}$, as described in [**264**]. We return at the end of Chapter 2 of this book to discuss the Archimedean local factors of the L-functions of motives from the point of view of noncommutative geometry.

8.1.4. *The Weil conjectures.* An aspect of the theory of zeta functions of algebraic varieties that had a very important historical role in the development of the idea of motives was a set of conjectures on the properties of the zeta functions of (1.374) and (1.377), partly modeled on the Riemann zeta function. These are the Weil conjectures, which were initially proved by Weil in the case of curves over finite fields and later in complete generality, for any nonsingular projective variety over a finite field, by Deligne in [**110**], after important contributions of Grothendieck, Dwork, Artin, that obtained a proof of the first and second of the Weil conjectures.

The first conjecture predicts that the zeta function $Z_X(t)$ is a rational function of t, so that one can write

$$Z_X(t) = \prod_{\ell=0}^{2\dim X} P_\ell(t),$$

with the $P_\ell(t)$ polynomials of the form

$$P_\ell(t) = \prod_j (1 - \lambda_{\ell j} t), \tag{1.382}$$

with $\lambda_{\ell j}$ the eigenvalues of the Frobenius on $H^\ell(X)$ as in (1.377). The second conjecture predicts a functional equation for $Z_X(t)$, where for $t = q^{-s}$ one has

$$Z_X(q^{(-\dim X + s)}) = \pm q^{\chi(X)(-s+\dim X/2)} Z_X(q^{-s}), \tag{1.383}$$

with $\chi(X) = \Delta \bullet \Delta$ the self intersection of the diagonal, i.e. the Euler characteristic. Finally, the third conjecture, which is the most difficult, is the analog of the Riemann Hypothesis and gives a constraint on the position of the zeros of the factors $P_\ell(t)$. Namely, the conjecture predicts that, setting $t = q^{-s}$, the zeros of P_ℓ lie on the line $\Re(s) = \ell/2$. This can be equivalently stated as the condition that the eigenvalues $\lambda_{\ell j}$ are of absolute value

$$|\lambda_{\ell j}| = q^{\ell/2}. \tag{1.384}$$

8.1.5. *Jacobians.* Another fundamental ingredient that played a role in the development of the theory of motives is the geometry of algebraic curves and their Jacobians. It was again André Weil who proposed to use the Jacobian $J(C)$ of a smooth projective curve C as a geometric replacement for the cohomology $H^1(C, \mathbb{Q})$. In fact, prior to the general development of étale cohomology, the fact that the Jacobian admits a purely algebraic construction provided one with a good cohomology theory for curves over any field of definition. This was especially useful in studying the properties of the zeta function of a curve over a finite field, which is the direct analog of the Riemann zeta function in the case of a function field (the latter being the field of functions of a smooth projective curve over a finite field).
In motivic terms this property of the Jacobian is expressed by the identification

$$\mathrm{Hom}_{\mathcal{M}^{\mathrm{eff}}_{\mathrm{num}}(k)_{\mathbb{Q}}}(h^1(C), h^1(C')) = \mathrm{Hom}(J(C), J(C')) \otimes \mathbb{Q}. \tag{1.385}$$

This means that it makes sense to think of $J(C)$ as the motive $h^1(C)$, so that the additive decomposition of a smooth projective curve in terms of motives takes the form

$$h(C) = 1 \oplus J(C) \oplus L.$$

In higher dimension, for a smooth projective variety X, a good generalization of the Jacobian is provided by the Albanese and Picard varieties $\mathrm{Alb}(X)$ and $\mathrm{Pic}(X)$. These also admit a purely algebraic construction, they are dual abelian varieties and both are isomorphic to the Jacobian in the case of curves (see e.g. [**248**] for a survey of their role in the theory of motives).

8.1.6. *Galois theory.* Classical Galois theory has also been absorbed into the theory of motives, as the particular case of Artin motives, which we discuss in more detail below. In this case, the main idea is that the theory of motives gives a "linearization" of the theory of varieties. In particular, the case of Galois theory arises by considering the case of zero-dimensional smooth projective varieties (i.e. finite sets of points). The Grothendieck–Galois correspondence states that there is an equivalence of categories between the category $\mathcal{V}^0(k)$ of zero-dimensional smooth projective varieties over k and the category of finite sets with a continuous action of $\mathrm{Gal}(k^{\mathrm{sep}}/k)$, with k^{sep} a separable closure of k. The functor that realizes the equivalence of categories is given by taking the set of algebraic points of a variety

$$X \mapsto X(k^{\mathrm{sep}}).$$

The theory of Artin motives provides a linearization of the Grothendieck–Galois correspondence in the following way. One obtains an equivalence of categories between the category $\mathcal{AM}_k$ (also denoted by $\mathcal{M}_{Artin}(k)$) of Artin motives over k with $\mathbb{Q}$-coefficients and the category of finite-dimensional $\mathbb{Q}$-linear representations Rep_G of $G = \mathrm{Gal}(k^{\mathrm{sep}}/k)$. The category $\mathcal{AM}_k$ is the subcategory of $\mathcal{M}^{\mathrm{eff}}_{\mathrm{num}}(k)$ generated by the zero-dimensional objects.
In this way Galois theory becomes the first instance of an application of the Tannakian formalism to motives, with the Galois group $\mathrm{Gal}(k^{\mathrm{sep}}/k)$ as the *motivic Galois group* of the Tannakian category of Artin motives. In this case the functor that realizes the equivalence of categories is the fiber functor

$$X \mapsto \mathbb{Q}^{X(k^{\mathrm{sep}})}.$$

8.1.7. *Modular forms.* The theory of motives has important connections with automorphic forms, most notably through some aspects of the Langlands program. We are not going to touch on this aspect at all, and in fact we will only encounter elliptic modular forms in Chapter 3. It is at least worth mentioning that one can associate motives to elliptic modular forms (i.e. modular forms for congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$), using the geometry of Kuga–Sato varieties or of the moduli space of stable curves of genus one with marked points (these constructions were obtained, respectively, in [**260**] and [**101**]).

8.2. Pure motives.

The history of motives began in the early sixties, when Grothendieck envisioned the whole idea of the existence of a theory of motives underlying the different cohomology theories of algebraic varieties. Although Grothendieck himself did not write about it at the time, the first important paper describing the fundamental steps in the construction of a category of pure motives and its properties appeared shortly afterwards, written by Manin [**213**]. This seminal paper contained the first new result in algebraic geometry obtained as a direct consequence of the theory of motives, namely a proof of the Weil conjecture for the zeta function of a unirational projective 3-fold over a finite field. Even though eventually Deligne's proof of the Weil conjectures in full generality did not use directly the theory of motives, this result provided convincing evidence for the original vision of Grothendieck that the

motivic formalism was to provide the deep rationale underlying the Weil conjectures as well as a wider set of conjectural properties of algebraic cycles, still largely unexplored today, collectively known as the "standard conjectures" (see below).
After this initial development, the progress of the theory of motives was somewhat slow. The crucial difficulty lies in the fact that not enough is known in general about algebraic cycles in smooth projective varieties.
For the theory of pure motives, we refer the interested reader to the very nice survey by Serre [**263**], as well as to the references [**2**], [**108**], [**155**], [**174**], [**186**], [**213**], [**235**].

8.2.1. *A Tannakian category.*

We have already mentioned the category $\mathcal{M}^{\rm eff}_{\sim}(k)$ for various possible equivalence relations on algebraic cycles. In particular, $\mathcal{M}^{\rm eff}_{\rm num}(k)$ is by [**174**] an abelian category. The monoidal structure induced by the product of varieties can be improved to the structure of a rigid tensor category, if one formally inverts the Lefschetz motive L. One introduces in this way additional objects, the Tate motives, with $\mathbb{Q}(1) = L^{-1}$ and $\mathbb{Q}(n+m) \simeq \mathbb{Q}(n) \otimes \mathbb{Q}(m)$. This makes it possible to define the duality which is part of the rigid tensor structure.
More precisely, one enlarges the category $\mathcal{M}^{\rm eff}_{\rm num}(k)$ to a category $\mathcal{M}_{\rm num}(k)$ where the objects are of the form (X,p,n), with (X,p) as before and $n \in \mathbb{Z}$, and the morphisms are given by

$$\mathrm{Hom}_{\mathcal{M}_{\rm num}(k)}((X,p,n),(Y,q,m)) = q \circ C^{m-n}(X,Y)_{num,\mathbb{Q}} \circ p, \tag{1.386}$$

where $C^{m-n}(X,Y)_{\rm num,\mathbb{Q}}$ denotes the $\mathbb{Q}$-vector space of algebraic cycles in $X \times Y$ of codimension equal to $\dim X + m - n$ modulo numerical equivalence. One often writes an object $M = (X,p,n)$ with the notation $M = p\,h(X)(n)$, where (n) stands for the Tate twist $\otimes\mathbb{Q}(n)$. One finds the old $\mathcal{M}^{\rm eff}_{\rm num}(k)$ embedded as a subcategory with objects $(X,p,0)$, and the new objects are obtained by adding the Tate motives. The tensor product is of the form

$$(X,p,n) \otimes (Y,q,m) = (X \times Y, p \otimes q, n+m)$$

and the duality is given by

$$(X,p,n)^\vee = (X, p^t, \dim X - n),$$

with p^t the transpose. This satisfies the properties of a rigid tensor category. The subcategory $\mathcal{M}^{\rm eff}_{\rm num}(k)$ is not stable under duality.

$\mathcal{M}_{\rm num}(k)$ is called the category of pure motives over k, with the numerical equivalence.

One can define in a similar way categories $\mathcal{M}_{\sim}(k)$, for other equivalence relations, obtained from the corresponding $\mathcal{M}^{\rm eff}_{\sim}(k)$ by adding the Tate motives. However, the result of Jannsen [**174**] shows that only in the case of the numerical equivalence is $\mathcal{M}_{\sim}(k)$ a semi-simple abelian category.

Having a category of pure motives $\mathcal{M}_{\rm num}(k)$ that is a semi-simple abelian category with a rigid tensor structure, one can look for the existence of fiber functors that will make it into a Tannakian category.

A result of Deligne [**106**] shows that if $\mathcal{C}$ is an abelian category over a field K of characteristic zero with a rigid tensor structure and such that $\mathrm{End}(1) = K$, then $\mathcal{C}$ is Tannakian if and only if for all $X \in \mathrm{Obj}(\mathcal{C})$ the trace takes values $\mathrm{Tr}(1_X) \in \mathbb{N}$. This

immediately shows that the tensor structure described above has to be modified if one wants to obtain a Tannakian category. In fact, with the above structure one would obtain $\mathrm{Tr}(1_X) = \chi(X) \in \mathbb{Z}$. This can be remedied by changing the tensor product structure, by replacing the commutativity axiom with a graded commutative version, by which the natural isomorphism of $M \otimes N$ and $N \otimes N$ is twisted by a sign $(-1)^{ij}$, for M and N respectively of weight i and j.
The standard conjectures (of which more below) imply that $\mathcal{M}_{\mathrm{num}}(k) = \mathcal{M}_{\mathrm{hom}}(k)$. With the modification to the tensor product structure mentioned above, one obtains that this category is a Tannakian category. The realization to the Weil cohomology H^*, say Betti cohomology, gives a fiber functor

$$\omega : \mathcal{M}_{\mathrm{hom}}(k) \to \mathrm{Vect}_{\mathbb{Q}}$$

which on $M = h^i(X)$ gives $\omega(M) = H^i_{\mathrm{B}}(X, \mathbb{Q})$. Tensoring with a Tate motive $\mathbb{Q}(n)$ translates into the Tate twist by $(2\pi i)^n$.
Thus, provided the standard conjectures hold, the category $\mathcal{M}_{\mathrm{num}}(k) = \mathcal{M}_{\mathrm{hom}}(k)$ is a neutral Tannakian category. We write just $\mathcal{M}(k)$, as we are going to restrict to just this equivalence relation. Since it is semi-simple, the identity component of its motivic Galois group $G_{\mathrm{mot}}(\mathcal{M}(k))$ is a pro-reductive group. Let $\mathcal{C} \subset \mathcal{M}(k)$ be a full subcategory which is stable under tensor product, duality, and such that subobjects or quotients of objects of $\mathcal{C}$ are objects of $\mathcal{C}$. Then $\mathcal{C}$ inherits the structure of Tannakian category and there is a corresponding morphism

$$G_{\mathrm{mot}}(\mathcal{M}(k)) \to G_{\mathrm{mot}}(\mathcal{C}). \tag{1.387}$$

For instance, one can consider the Tannakian subcategory $\mathcal{TM}(k)$ of $\mathcal{M}(k)$ generated by the Tate motives. This is the category of pure Tate motives. Its motivic Galois group is just the multiplicative group $\mathbb{G}_m$. This gives an induced morphism

$$t : G_{\mathrm{mot}}(\mathcal{M}(k)) \to \mathbb{G}_m. \tag{1.388}$$

The fact that the realization H^* is compatible with the $\mathbb{Z}$-grading also gives a morphism $w : \mathbb{G}_m \to G_{\mathrm{mot}}(\mathcal{M}(k))$ with the property that $t \circ w = -2 \in \mathrm{End}(\mathbb{G}_m) = \mathbb{Z}$, since $\mathbb{Q}(1)$ is of weight -2, being the inverse of $h^2(\mathbb{P}^1)$.
A more general example of the morphism (1.387) is obtained by considering the subcategory $\mathcal{M}(M)$ of $\mathcal{M}$ generated by a single object M. Then the identity component of the reductive group $G_{\mathrm{mot}}(\mathcal{M}(M))$ is the Mumford–Tate group of M. For example, suppose that $M = (E, 1, 0)$ is an elliptic curve E. Then in the case without complex multiplication the Mumford–Tate group is GL_2 while if E has complex multiplication by an imaginary quadratic field $\mathbb{K}$ then the Mumford–Tate group is the algebraic torus given by viewing $\mathbb{K}^*$ as an algebraic torus over $\mathbb{Q}$. The case of Tate motives corresponds to the fact that the Mumford–Tate group of $\mathbb{Q}(1)$ is $\mathbb{G}_m$. For more information on motivic Galois groups, we refer the reader to [**264**], [**2**].

8.2.2. *Artin motives.* Another important example of a Tannakian subcategory of the category $\mathcal{M}(k)$ is the category $\mathcal{AM}(k)$ of Artin motives. This is the subcategory generated by objects $M = (X, p, 0)$, where X is a zero-dimensional algebraic variety.

Since one is only dealing with zero-dimensional varieties, all the equivalence relations on algebraic cycles in the product are trivial, so that it does not matter in which $\mathcal{M}_\sim(k)$ one views Artin motives. Moreover, one does not need to add the Tate

motives to have duality, nor does one need any modification in the commutativity axiom of the tensor product.
Artin motives form an abelian category, and in fact a neutral Tannakian category, with the fiber functor

$$\omega(X) = \mathbb{Q}^{X(k^{\rm sep})} = H^0_{\rm B}(X, \mathbb{Q}).$$

As we mentioned above the motivic Galois group is in this case

$$G_{\rm mot}(\mathcal{AM}(k)) = {\rm Gal}(k^{\rm sep}/k), \tag{1.389}$$

with a morphism

$$a : G_{\rm mot}(\mathcal{M}(k)) \to {\rm Gal}(k^{\rm sep}/k). \tag{1.390}$$

As discussed above, Artin motives provide a linear version of the correspondence of Grothendieck–Galois and the identification (1.389) gives an interpretation of motivic Galois groups as a powerful generalization of ordinary Galois groups.

8.2.3. *Motives with coefficients.* We have so far considered the category $\mathcal{M}_\sim(k)$ of pure motives as a $\mathbb{Q}$-linear category. One can also consider categories of pure motives with coefficients in a field $\mathbb{E}$ of characteristic zero, which we denote by $\mathcal{M}_\sim(k)_{\mathbb{E}}$. This can be described as the (pseudo)abelian envelope of the additive category generated by ${\rm Obj}(\mathcal{M}_\sim(k))$ with morphisms

$${\rm Hom}_{\mathcal{M}_\sim(k)_{\mathbb{E}}}(M, N) = {\rm Hom}_{\mathcal{M}_\sim(k)}(M, N) \otimes \mathbb{E}. \tag{1.391}$$

Notice that, having modified the morphisms, one in general will have to add new objects to still have a (pseudo)abelian category.
Notice that the notion of motives with coefficients that we use here is compatible with another, apparently different, notion which is often adopted in the literature. In fact, in the case where the field of definition k of the motives is of characteristic zero and $\mathbb{E}$ is a *finite* extension of k, one can say (cf. [**282**] and [**109**] §2.1) that a motive M has coefficients in $\mathbb{E}$ if there is a homomorphism $\mathbb{E} \to {\rm End}(M)$. This description is compatible with defining, as we did above, $\mathcal{M}_\sim(k)_{\mathbb{E}}$ as the smallest (pseudo)abelian category containing the pure motives over k with morphisms (1.391). This is shown in [**109**] §2.1: first notice that, for k of characteristic zero and $\mathbb{E}$ a finite extension of k, one can associate to a motive M over k an element of $\mathcal{M}_\sim(k)_{\mathbb{E}}$ by taking $M \otimes_k \mathbb{E}$, with the obvious $\mathbb{E}$-module structure. If M is endowed with a map $\mathbb{E} \to {\rm End}(M)$, then the corresponding object in $\mathcal{M}_\sim(k)_{\mathbb{E}}$ is the range of the projector onto the largest direct summand on which the two structures of $\mathbb{E}$-module agree.
Notice that, whereas for a finite extension $\mathbb{E}$ one has in this way two notions of "motives with coefficients" that agree by the argument above, only the definition of $\mathcal{M}_\sim(k)_{\mathbb{E}}$ as the smallest (pseudo)abelian category containing the pure motives over k with morphisms (1.391) continues to make sense when $\mathbb{E}$ is an arbitrary field of characteristic zero. In particular, it can be used to define $\mathcal{M}_\sim(k)_{\mathbb{C}}$. This observation will be useful in the last chapter of this book, when we discuss an interplay between Artin motives and noncommutative spaces described by C^*-algebras.

8.2.4. *Grothendieck's standard conjectures.* As a way to approach the Weil conjectures, Grothendieck formulated in [**155**] a set of conjectures on algebraic cycles, called the "standard conjectures", which are basic to the theory of motives and imply the Weil conjectures (see [**155**], [**185**]). While the latter were proved by Deligne using a different method, the standard conjectures remain unproven, except for special cases. In subsequent developments of the theory of motives people adapted to work conditionally or to establish unconditional results that bypass the standard conjectures, but this set of conjectures remains a fundamental unsolved problem in the theory.

- *Lefschetz type:* Given a Weil cohomology H^*, the Lefschetz isomorphisms (1.370) determine a Lefschetz involution on $\oplus_{i,r} H^i(X)(r)$ given by $\mathcal{L}^{\dim X - i}$ for $i \leq \dim X$ and $\mathcal{L}^{i - \dim X}$ for $i > \dim X$. The conjecture states that this involution is realized by an algebraic cycle with $\mathbb{Q}$-coefficients.
- *Hodge type:* The Hodge involution is obtained by modifying the Lefschetz involution by a factor on each primitive component of the cohomology (cf. e.g. [**2**] §5). The conjecture states that the Hodge involution determines a positive definite quadratic form on the space $C(X)_{\mathrm{hom},\mathbb{Q}}$ of algebraic cycles, with values in $\mathbb{Q}$.
- *Künneth type:* The projectors $H^*(X) \to H^i(X)$ are realized by algebraic cycles, i.e. by idempotents $p^i \in C^0(X,X)_{\mathrm{hom},\mathbb{Q}}$.
- $\sim_{\mathrm{num}} = \sim_{\mathrm{hom}}$*:* Numerical and homological equivalences coincide.

For varieties over a field k of characteristic zero, the first conjecture (of Lefschetz type) implies all the others. In general, Lefschetz implies Künneth and Lefschetz with Hodge imply all the others. A lot is known in specific cases, for example in the case of abelian varieties [**185**].

8.3. Mixed motives.

When, along with smooth projective varieties, one wants to consider more general classes of varieties and schemes, one leaves the theory of pure motives for the much more mysterious world of *mixed motives*. At first, and especially thinking in terms of zeta functions, it would seem that one can treat more general varieties by still working with pure motives but considering the associated Grothendieck group. As we discuss below, this is in fact not sufficient as it does not take into account a very important feature of mixed motives: the presence of nontrivial extensions of pure motives. This can be already appreciated by looking at the most concrete example of mixed motives, which is provided by the 1-motives. The problem of constructing categories of mixed motives, which continue to have the right properties of a "universal cohomology theory" for algebraic varieties, is an extremely challenging task. Recent results such as [**287**], [**208**], [**162**] produced good categories of mixed motives, but these are still in general very difficult objects to deal with. We refer the reader to Bloch's lectures [**21**] for an overview of mixed motives.

8.3.1. *The Grothendieck group of varieties.* Given an algebraic variety X over a field of characteristic zero (where Hironaka's resolution of singularities holds), one can always write such X as a disjoint union of $X_i - D_i$, where the X_i are smooth projective and the D_i are lower dimensional. Thus, inductively on the dimension, one can assign to X a virtual object $[X]$ in a Grothendieck group of algebraic varieties $[X] = \sum_i [Y_i] - \sum_j [Z_j]$ where the Y_i and Z_j are smooth projective varieties.
More precisely, the Grothendieck group $K_0(\mathcal{V}(k))$ of varieties over k is the quotient of the free abelian group generated by symbols $[X]$ by the scissor-congruence relation

$$[X] = [Y] + [X \setminus Y], \tag{1.392}$$

for any closed subvariety $Y \subset X$, and the isomorphism relation

$$[X] = [Y], \quad \text{if } X \simeq Y \text{ are } k\text{-isomorphic.} \tag{1.393}$$

It has a product given by $[X][Y] = [X \times_k Y]$. One writes $L = [\mathbb{A}^1]$ for the class of the affine line. This corresponds to the Lefschetz motive.

It is also possible to consider the Grothendieck group of the category of pure motives $\mathcal{M}_\sim(k)$. The Grothendieck group $K_0(\mathcal{C})$ of an abelian category $\mathcal{C}$ is defined by first considering the free abelian group generated by the isomorphism classes $[M]$ of objects $M \in \mathrm{Obj}(\mathcal{C})$ (assuming this is a set), and then imposing the following relations. For any exact sequence $0 \to M \to M' \to M'' \to 0$ one has a relation $[M'] = [M] + [M'']$.

There is a homomorphism $K_0(\mathcal{V}(k)) \to K_0(\mathcal{M}_\sim(k))$, which extends to a homomorphism $K_0(\mathcal{V}(k))[L^{-1}] \to K_0(\mathcal{M}_\sim(k))$. The image of the class $[X]$ of a smooth projective variety X is the class of the pure motive $h(X) = (X, 1, 0)$. The image of $L = [\mathbb{A}^1]$ is the class of the Lefschetz motive.

The zeta functions can be lifted to the level of the Grothendieck group. For example, in the case of a finite field $k = \mathbb{F}_q$ one can write (1.374) equivalently as

$$Z_X(t) = \exp\left(\sum_n \frac{\#X(\mathbb{F}_{q^n})}{n}\, t^n\right) = \sum_{n\geq 0} \#s^n(X)(\mathbb{F}_q)\, t^n \tag{1.394}$$

where $s^n(X)$ denotes the n-th symmetric power of X. This leads to Kapranov's formulation of a zeta function

$$Z_X(t) = \sum_{n\geq 0} [s^n(X)]\, t^n \tag{1.395}$$

for an arbitrary field k, with values in the Grothendieck group from which the ordinary one is obtained in the case $k = \mathbb{F}_q$ by applying a "counting of points" homomorphism from $K_0(\mathcal{V}(k))$. One has $Z_{\mathbb{P}^1}(t) = (1-t)^{-1}(1-[\mathbb{A}^1]t)^{-1}$. Kontsevich's theory of motivic integration also lives naturally in (a completion of) $K_0(\mathcal{V}(k))[L^{-1}]$.

While working with "virtual pure motives" in the Grothendieck group suffices for some purposes, if one wants a theory that satisfies the main requirements of a category of motives, including providing a universal cohomology theory (via the Ext functors), working with the Grothendieck group is not sufficient.

8.3.2. *1-motives.* A first concrete example of mixed motives is given by the 1-motives. Deligne introduced 1-motives in [**111**] as geometries underlying certain classes of mixed Hodge structures. They can be thought of as a special case of mixed motives, just as Jacobians and abelian varieties provide a good class of pure motives. They exhibit the main properties one expects for mixed motives, namely: they have a weight filtration whose graded pieces are pure motives and, when considered over $\mathbb{C}$, they give rise to mixed Hodge structures.

A semi-abelian variety is an extension (in the category of algebraic groups) of an abelian variety A by an algebraic torus T. While abelian varieties are projective–hence they live in the category of pure motives–algebraic tori are affine varieties. Moreover, the fact that we are considering non-trivial extensions reveals that one is leaving the world of pure motives for mixed motives. A 1-motive M over k is a complex of the form $[\Lambda \xrightarrow{u} G]$, where Λ is a finitely generated torsion free abelian group and G is a semi-abelian variety. The map u is a homomorphism $u : \Lambda \to$

$G(k)$. A morphism of 1-motives M, M' is a commutative diagram relating the two complexes defining M and M'. One needs to invert isogenies to have an abelian category. Duality on 1-motives combines Pontrjagin duality and duality of abelian varieties. The weight filtration on a 1-motive M has $W_{-3}M = 0$, $W_{-2}M = [0 \to T]$, $W_{-1} = [0 \to G]$, $W_0 = M$ with respective quotients T, A, and Λ. The category of 1-motives contains the category of Artin motives: these are identified with objects $M = [\Lambda \to 0]$.
Deligne conjectured in [**111**] that there are purely algebraic constructions for certain 1-motives arising from the mixed Hodge structure on $H^n(X, \mathbb{Z}(1))$ of a complex algebraic variety X. This generalizes the fact that Jacobians, Albanese and Picard varieties admit algebraic constructions. The conjecture was recently proved by Ramachandran in [**249**].

8.3.3. *Categories of mixed motives.* More generally, instead of working over a field k, one can consider a category $\mathcal{M}_{\rm mix}(S)$ of motives (or "motivic sheaves") over a scheme S. In this case, the functors above are natural in S and to a map of schemes $f : S_1 \to S_2$ there correspond functors f^*, f_*, $f^!$, $f_!$, behaving like the corresponding functors of sheaves.
At present there is not yet a general construction of such a category of mixed motives $\mathcal{M}_{\rm mix}(S)$. There are, however, constructions of a triangulated tensor category $\mathcal{DM}(S)$, which has the right properties to be the bounded derived category of the category of mixed motives. The constructions of $\mathcal{DM}(S)$ of [**208**] and [**287**] are known to be equivalent. Other constructions are given in [**162**] and in unpublished work by Nori.
In general, given a construction of a triangulated tensor category, one can extract from it an abelian category by considering the *heart of a t-structure.* A caveat with this procedure is that it is not always the case that the given triangulated tensor category is in fact the bounded derived category of the heart of a t-structure.
The available constructions, in any case, are obtained via this general procedure of t-structures which can be summarized as follows. A triangulated category $\mathcal{D}$ is an additive category with an automorphism T and a family of distinguished triangles $X \to Y \to Z \to T(X)$, satisfying suitable axioms (which we do not recall here). We use the notation $\mathcal{D}^{\geq n} = \mathcal{D}^{\geq 0}[-n]$ and $\mathcal{D}^{\leq n} = \mathcal{D}^{\leq 0}[-n]$, with $X[n] = T^n(X)$ and $f[n] = T^n(f)$. A t-structure consists of two full subcategories $\mathcal{D}^{\leq 0}$ and $\mathcal{D}^{\geq 0}$ with the properties: $\mathcal{D}^{\leq -1} \subset \mathcal{D}^{\leq 0}$ and $\mathcal{D}^{\geq 1} \subset \mathcal{D}^{\geq 0}$; for all $X \in \mathcal{D}^{\leq 0}$ and all $Y \in \mathcal{D}^{\geq 1}$ one has $\mathrm{Hom}_{\mathcal{D}}(X, Y) = 0$; for all $Y \in \mathcal{D}$ there exists a distinguished triangle as above with $X \in \mathcal{D}^{\leq 0}$ and $Z \in \mathcal{D}^{\geq 1}$. The heart of the t-structure is the full subcategory $\mathcal{D}^0 = \mathcal{D}^{\leq 0} \cap \mathcal{D}^{\geq 0}$. It is an abelian category.

We have seen that, in the theory of pure motives, one can associate to a smooth projective algebraic variety X over a field k a motivic decomposition $\oplus_i h^i(X)$ (assuming the standard conjecture of Künneth type), where the $h^i(X)$ are objects of $\mathcal{M}_\sim(k)$. In this way, one views pure motives as a universal cohomology theory for smooth projective algebraic varieties. Moreover, the motivic cohomology $h(X) = \oplus_i h^i(X)$ has the property that, for any given Weil cohomology theory H^* there is a realization functor R_H satisfying $H^i(X) = R_H\, h^i(X)$.
On the category $\mathcal{M}_{\rm mix}(k)$ of mixed motives one expects to have *absolute cohomology* groups (cf. [**108**]) given by

$$H^i_{\rm abs}(M) = \mathrm{Ext}^i_{\mathcal{DM}(S)}(1, M). \tag{1.396}$$

for M a motive. Here the Ext^i are taken in the triangulated category $\mathcal{DM}(S)$ whose heart is the category $\mathcal{M}_{\mathrm{mix}}(S)$ of mixed motives. For a variety X the absolute cohomology would then be obtained from the motivic cohomology via a spectral sequence with

$$E_2^{pq} = H^p_{\mathrm{abs}}(h^q(X)) \Rightarrow H^{p+q}_{\mathrm{abs}}(X). \tag{1.397}$$

8.3.4. *Mixed Tate motives.* The triangulated category of *mixed Tate motives* $\mathcal{DTM}(S)$ is then defined as the full triangulated subcategory of $\mathcal{DM}(S)$ generated by the Tate objects. It is possible to define on it a t-structure whose heart gives a category of mixed Tate motives $\mathcal{TM}_{\mathrm{mix}}(S)$, provided the Beilinson–Soulé vanishing conjecture holds, namely when

$$\mathrm{Hom}^j(\mathbb{Q}(0), \mathbb{Q}(n)) = 0, \qquad \text{for } n > 0, j \leq 0. \tag{1.398}$$

where $\mathrm{Hom}^j(M, N) = \mathrm{Hom}(M, N[j])$. The conjecture (1.398) is known to hold in the case of $k = \mathbb{K}$ a number field (see [**207**]), where one has

$$\mathrm{Ext}^1_{\mathcal{DM}(\mathbb{K})}(\mathbb{Q}(0), \mathbb{Q}(n)) = K_{2n-1}(\mathbb{K}) \otimes \mathbb{Q} \tag{1.399}$$

and $\mathrm{Ext}^2_{\mathcal{DM}(\mathbb{K})}(\mathbb{Q}(0), \mathbb{Q}(n)) = 0$. Here $K_{2n-1}(\mathbb{K})$ is Quillen's higher algebraic K-theory (cf. [**208**] Appendix B). Thus, in this case it is possible to extract from the triangulated tensor category a Tannakian category $\mathcal{TM}_{\mathrm{mix}}(\mathbb{K})$ of mixed Tate motives, with fiber functor ω to $\mathbb{Z}$-graded $\mathbb{Q}$-vector spaces, $M \mapsto \omega(M) = \oplus_n \omega_n(M)$ with

$$\omega_n(M) = \mathrm{Hom}(\mathbb{Q}(n), \mathrm{Gr}^w_{-2n}(M)), \tag{1.400}$$

where $\mathrm{Gr}^w_{-2n}(M) = W_{-2n}(M)/W_{-2(n+1)}(M)$ is the graded structure associated to the finite increasing weight filtration W.

As an example, the mixed Tate motives that correspond to extensions in

$$\mathrm{Ext}^1_{\mathcal{DM}(\mathbb{K})}(\mathbb{Q}(0), \mathbb{Q}(1))$$

are the Kümmer motives. These are a particular case of 1-motives, of the form $M = [\mathbb{Z} \xrightarrow{u} \mathbb{G}_m]$, with $u(1) = q \in \mathbb{K}^*$.

The motivic Galois group of the category $\mathcal{TM}_{\mathrm{mix}}(\mathbb{K})$ is then an extension $G = U \rtimes \mathbb{G}_m$, where the reductive piece is $\mathbb{G}_m$ as in the case of pure Tate motives, while U is prounipotent. By the results of Goncharov and Deligne–Goncharov (see [**146**], [**113**]), it is known that the pro-unipotent affine group scheme U corresponds to a graded Lie algebra Lie (U) that is free with one generator in each odd degree $n \leq -3$.

8.3.5. *Mixed Tate motives and the "cosmic Galois group".* A description of the category $\mathcal{TM}_{\mathrm{mix}}(S)$ in terms of algebraic K-theory is also possible in the case where the scheme S is the set of V-integers $\mathcal{O}_V$ of a number field $\mathbb{K}$, for V a set of finite places of $\mathbb{K}$. In this case, objects of $\mathcal{TM}_{\mathrm{mix}}(\mathcal{O}_V)$ are mixed Tate motives over $\mathbb{K}$ that are unramified at each finite place $v \notin V$. For $\mathcal{TM}_{\mathrm{mix}}(\mathcal{O}_V)$ we have

$$\mathrm{Ext}^1_{\mathcal{DM}(\mathcal{O}_V)}(\mathbb{Q}(0), \mathbb{Q}(n)) = \begin{cases} K_{2n-1}(\mathbb{K}) \otimes \mathbb{Q} & n \geq 2 \\ \mathcal{O}_V^* \otimes \mathbb{Q} & n = 1 \\ 0 & n \leq 0. \end{cases} \tag{1.401}$$

and $\mathrm{Ext}^2_{\mathcal{DM}(\mathcal{O}_V)}(\mathbb{Q}(0),\mathbb{Q}(n)) = 0$. In fact, the difference between the Ext in $\mathcal{TM}_{\mathrm{mix}}(\mathcal{O}_V)$ of (1.401) and the Ext in $\mathcal{TM}_{\mathrm{mix}}(\mathbb{K})$ of (1.399) is

$$\mathrm{Ext}^1(\mathbb{Q}(0),\mathbb{Q}(1)),$$

which is finite-dimensional in (1.401) and infinite-dimensional in (1.399). The category $\mathcal{TM}_{\mathrm{mix}}(\mathcal{O}_V)$ is also a neutral Tannakian category, and the fiber functor determines an equivalence of categories between $\mathcal{TM}_{\mathrm{mix}}(\mathcal{O}_V)$ and finite-dimensional linear representations of an affine group scheme of the form $U \rtimes \mathbb{G}_m$ with U prounipotent. The Lie algebra $\mathrm{Lie}(U)$ is freely generated by a set of homogeneous generators in degree n identified with a basis of the dual of $\mathrm{Ext}^1(\mathbb{Q}(0),\mathbb{Q}(n))$ (cf. Prop. 2.3 of [**113**]). There is however no *canonical* identification between $\mathrm{Lie}(U)$ and the free Lie algebra generated by the graded vector space $\oplus\mathrm{Ext}^1(\mathbb{Q}(0),\mathbb{Q}(n))^\vee$.

We mention the following case, which will be the one most relevant in the context of perturbative renormalization.

PROPOSITION 1.110. ([**113**], [**146**]) *Consider the scheme $S_N = \mathrm{Spec}\,\mathcal{O}[1/N]$ for $k = \mathbb{Q}(\zeta_N)$ the cyclotomic field of level N and $\mathcal{O}$ its ring of integers. For $N = 3$ or 4, the motivic Galois group of the category $\mathcal{TM}_{\mathrm{mix}}(S_N)$ is of the form $U \rtimes \mathbb{G}_m$, where the Lie algebra $\mathrm{Lie}(U)$ is (noncanonically) isomorphic to the free graded Lie algebra with one generator e_n in each degree $n \leq -1$.*

The reason why this result is relevant to our discussion of renormalization is the following result, which immediately follows from the description of the “cosmic Galois group” $\mathbb{U}^*$.

COROLLARY 1.111. *There is an equivalence of categories between the neutral Tannakian category of flat equisingular vector bundles and $\mathcal{TM}_{\mathrm{mix}}(S_N)$ for $N = 3$ or 4.*

REMARK 1.112. Notice that in the case of mixed Tate motives, the grading is induced by the weight filtration on the mixed Tate motives. The latter is indexed by *even* integers. Thus, for a mixed Tate motive M one has $\omega_n(M)$ as in (1.400). The free graded Lie algebra generated by the elements of $\mathrm{Ext}^1(\mathbb{Q}(0),\mathbb{Q}(1))$ is obtained by considering the classes of the extensions of Gr^w_{-2n} by $\mathrm{Gr}^w_{-2(n+1)}$, as in [**113**], p.5. Thus, we are in fact working with Lie algebras and commutative Hopf algebras (even grading) and not with corresponding graded-commutative objects. This should be compared with our choice of even grading for the Hopf algebra of Feynman graphs, Remark 1.24.

8.4. Mixed Hodge structures.

Another important aspect of the theory of motives is its relation to Hodge theory. Motives are often constructed as geometries underlying certain Hodge structures (see e.g. the case of 1-motives mentioned above). Pure and mixed Hodge structures naturally arise in many contexts: on the cohomology of algebraic varieties, on rational homotopy groups, on the complexes of vanishing cycles. We recall here some general facts about mixed Hodge structure and we point out a structural analogy between the resulting Tannakian formalism and the one we encountered in the context of equisingular connections arising from perturbative renormalization. This analogy will be useful in Chapter 2, when we discuss the Archimedean local factors of the completed Hasse–Weil zeta function in the context of noncommutative geometry.

We recall some facts about mixed Hodge structures, following [**273**]. A $\mathbb{Q}$-mixed Hodge structure ($\mathbb{Q}$-MHS) consists of the data $M = (V, W_\bullet, F^\bullet)$ of a finite-dimensional $\mathbb{Q}$-vector space, with an increasing filtration $W.V$ (weight filtration) and a decreasing filtration $F^{\cdot}V_{\mathbb{C}}$ for $V_C = V \otimes_{\mathbb{Q}} \mathbb{C}$ (the Hodge filtration), satisfying

$$\mathrm{gr}^W_n V_{\mathbb{C}} = F^p \mathrm{gr}^W_n V_{\mathbb{C}} \oplus \overline{F^{n-p+1}} \mathrm{gr}^W_n V_{\mathbb{C}}, \tag{1.402}$$

where $\bar{F}^{\cdot}$ is the complex conjugate of $F^{\cdot}$ with respect to $V_{\mathbb{R}} \subset V_{\mathbb{C}}$. The condition (1.402) is equivalent to a Hodge decomposition

$$\mathrm{gr}^W_n V_{\mathbb{C}} = \oplus_{p+q=n} V^{p,q}, \tag{1.403}$$

with

$$V^{p,q} = F^p \mathrm{gr}^W_{p+q} V_{\mathbb{C}} \cap \overline{F^q} \mathrm{gr}^W_{p+q} V_{\mathbb{C}}. \tag{1.404}$$

The Hodge structure $(V, W_\bullet, F^\bullet)$ is pure of weight n if $W_n = V$ and $W_{n-1} = 0$. An example of pure Hodge structure of weight $-2n$ is the Tate Hodge structure with $V = (2\pi i)^n \mathbb{Q}$.

A morphism $M_1 \to M_2$ of $\mathbb{Q}$-MHS is a linear map $\phi : V_1 \to V_2$ such that $\phi(W_k V_1) \subset W_k V_2$ and $\phi(F^p V_{1,\mathbb{C}}) \subset F^p V_{2,\mathbb{C}}$. Tensoring of $\mathbb{Q}$-MHS is also defined, by considering $V_1 \otimes_{\mathbb{Q}} V_2$ with filtrations $W_m(V_1 \otimes_{\mathbb{Q}} V_2) = \sum_{i+j=m} W_i(V_1) \otimes_{\mathbb{Q}} W_j(V_2)$ and $F^p(V_1 \otimes_{\mathbb{Q}} V_2) = \sum_{r+s=p} F^r(V_{1,\mathbb{C}}) \otimes_{\mathbb{C}} F^s(V_{2,\mathbb{C}})$. The Tate Hodge structure defines, by tensoring, a Tate twist on the category of $\mathbb{Q}$-MHS.

A $\mathbb{Q}$-mixed Hodge structure $M = (V, W_\bullet, F^\bullet)$ is *mixed Tate* if

$$\mathrm{gr}^W_{2p} M = \oplus \mathbb{Q}(-p), \quad \text{and} \quad \mathrm{gr}^W_{2p-1} M = 0.$$

If M is a mixed Hodge–Tate structure over $\mathbb{Q}$, then one has an associated period matrix obtained as follows (cf. §2 of [**147**]). One has an isomorphism $V_{\mathbb{C}} = \oplus_p F^p V_{\mathbb{C}} \cap W_{2p} V_{\mathbb{C}}$ and canonical maps $F^p V_{\mathbb{C}} \cap W_{2p} V_{\mathbb{C}} \stackrel{\simeq}{\to} \mathrm{gr}^W_{2p}(V_{\mathbb{Q}}) \otimes_{\mathbb{Q}} \mathbb{C}$. This gives an isomorphism $S_{HT} : \oplus_p \mathrm{gr}^W_{2p}(V_{\mathbb{Q}}) \otimes_{\mathbb{Q}} \mathbb{C} \to V_{\mathbb{C}}$. A splitting of the weight filtration also gives an isomorphism $S_W : \oplus_p \mathrm{gr}^W_{2p}(V_{\mathbb{Q}}) \otimes_{\mathbb{Q}} \mathbb{C} \to V_{\mathbb{C}}$. The composite $S_{HT}^{-1} \circ S_W$ is the period map. A choice of basis of $\mathrm{gr}^W_{2p} V_{\mathbb{Q}}$ determines the corresponding period matrix, whose entries are periods of rational algebraic differential forms over relative cycles on mixed Tate motives. Thus, the period matrix defined in this way gives the comparison isomorphism between Betti and de Rham cohomology in the mixed case.

For example, for the Kümmer motive $[\mathbb{Z} \stackrel{u}{\to} \mathbb{G}_m]$ with $u(1) = q$ one has a period matrix of the form

$$\begin{pmatrix} \frac{1}{2\pi i} & \frac{\log q}{2\pi i} \\ 0 & 1 \end{pmatrix}.$$

One can in fact specify a mixed Hodge–Tate structure by assigning the period matrix. An example is given by the families $M_k(z)$ (variations of mixed Hodge structures) associated to the polylogarithms

$$\mathrm{Li}_k(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^k}$$

with matrices of the form (cf. [**23**])

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ -\mathrm{Li}_1(z) & 2\pi i & 0 & 0 & 0 & \cdots \\ -\mathrm{Li}_2(z) & 2\pi i\,\log z & (2\pi i)^2 & 0 & 0 & \cdots \\ -\mathrm{Li}_3(z) & \frac{2\pi i\,(\log z)^2}{2!} & (2\pi i)^2\,\log z & (2\pi i)^3 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{pmatrix} \tag{1.405}$$

One can similarly define real mixed Hodge structures $\mathbb{R}$-MHS, by the same procedure, replacing $\mathbb{Q}$ with $\mathbb{R}$. More generally, one can consider mixed Hodge structures of the form $(A, W_\bullet, F^\bullet)$ with A an object in an abelian category, endowed with an increasing filtration $W.A$ and decreasing filtrations $F^\cdot A$ and $\bar{F}^\cdot A$ with $\mathrm{gr}_F^p\,\mathrm{gr}_{\bar{F}}^q\,\mathrm{gr}_n^W A = 0$ except when $n = p + q$.

Thus, one can consider the category of vector spaces over the field k endowed with filtrations $W_\bullet$, $F^\bullet$ and $\bar{F}^\bullet$ as above. We assume that k is of characteristic zero. Notice that, in this case, there are two natural choices of a fiber functor to vector spaces, one is the forgetful functor $\omega(V, W_\bullet, F^\bullet) = V$ and the other is $\omega_W(V, W_\bullet, F^\bullet) = \mathrm{gr}^W(V) = \oplus_n \mathrm{gr}_n^W(V)$. These two fiber functors are isomorphic through the identification $V \cong \mathrm{gr}^W(V)$ induced by the Hodge filtration $F^\bullet$. Namely, upon writing $\mathrm{gr}^W(V)$ as in (1.403) with $V^{p,q}$ as in (1.404), one can show that there is an isomorphism a_F between $(V, W_\bullet, F^\bullet)$ and $(\mathrm{gr}^W(V), W_\bullet, F_W^\bullet)$, induced by isomorphisms of the $V^{p,q}$ with

$$V_F^{p,q} = (W_{p+q} \cap F^p) \cap \left((W_{p+q} \cap \bar{F}^q) + \sum_{i \geq 0} (W_{p+q-i} \cap \bar{F}^{q-i+1}) \right).$$

This isomorphism a_F induces the isomorphism of the fiber functors. There is a similar isomorphism obtained using the opposite filtration $\bar{F}^\bullet$ and we write $e = a_{\bar{F}} a_F^{-1}$. This automorphism defines a 1-parameter subgroup of the motivic Galois group of the Tannakian category of mixed Hodge structures and, as we see below, it plays exactly the same role as the generator of the renormalization group in the case of perturbative renormalization.

Deligne shows in [**107**] that the category of k-vector spaces with filtrations $W_\bullet$, $F^\bullet$ and $\bar{F}^\bullet$ as above is a neutral Tannakian category. The Galois group of the Tannakian category is of the form

$$G_{\mathrm{MHS}} = \mathbb{U}_{\mathrm{MHS}} \rtimes (\mathbb{G}_m \times \mathbb{G}_m), \tag{1.406}$$

where $\mathbb{U}_{\mathrm{MHS}}$ is a pro-unipotent affine group scheme. This is dual to the commutative Hopf algebra $\mathcal{H}_{\mathbb{U}_{\mathrm{MHS}}} = U(\mathcal{L}_{\mathrm{MHS}})^\vee$, where the Lie algebra $\mathcal{L}_{\mathrm{MHS}}$ has generators $e_{-n,-m}$ of bi-degree (n, m), for each pair of integers $n, m \in \mathbb{N} = \mathbb{Z}_{>0}$. The action of $\mathbb{G}_m \times \mathbb{G}_m$ is related to the bigrading in the form

$$\mathbb{G}_m \times \mathbb{G}_m \ni (u, v) : \ell \mapsto u^{-n} v^{-m} \ell,$$

for $\ell \in \mathcal{L}_{\mathrm{MHS}}$ of bidegree (n, m). The element $d = a_{\bar{F}} a_F^{-1}$ is obtained as $d = \exp(e)$ with $e = \sum_{n,m} e_{-n,-m}$. This is clearly the analog of the generator of the renormalization group $e = \sum_n e_{-n}$ in the case of the "cosmic Galois group".

This applies in particular to the case of real mixed Hodge structures, though not to the case of rational mixed Hodge structures. In fact, for $\mathbb{Q}$-MHS the decreasing Hodge filtration is only a filtration of the complexified $V_{\mathbb{C}}$. In the case of $\mathbb{R}$-MHS the general result above specializes to give the Galois group $G_{\mathbb{R}-\mathrm{MHS}}$ as the real part of G_{MHS} with respect to the complex conjugation that maps $(u, v) \in \mathbb{G}_m \times \mathbb{G}_m$ to $(\bar{v}, \bar{u})$ and the generator $e_{-n,-m}$ to $-e_{-m,-n}$.

It is worth mentioning the fact that, for mixed Tate motives over a number field the functor to mixed Hodge structures is faithful and exact. This follows from Proposition 2.14 of [**113**], upon realizing $\mathcal{TM}_{\mathrm{mix}}(\mathbb{K})$ as a subcategory of Voevodsky's category [**287**]. The key point is the result, already mentioned in §§8.3.4 and 8.3.5 above, on the injectivity of the regulator map on the Beilinson–Soulé vanishing conjecture for K-theory, which is known to hold for number fields by Borel's computation of their K-groups, cf. [**207**].

8.5. Tate motives, periods, and quantum fields.

The fact that the identification of $\mathrm{Lie}(U)$ of Proposition 1.110 with the free graded Lie algebra $\mathrm{Lie}(\mathbb{U})$ with generators $\{e_n : n \leq -1\}$ is non-canonical means that the result of Corollary 1.111 only provides an abstract identification. One would like to have a much more concrete relation between the data of perturbative renormalization and mixed Tate motives.

In fact, there are strong reasons to believe that there should be such a relation between Feynman integrals and mixed Tate motives. We summarize a few facts, conjectures, and results that help in clarifying this general picture.

- The extensive computations of Broadhurst and Kreimer [**35**] show that multiple zeta values appear as residues of Feynman graphs.
- Multiple zeta values are periods of mixed Tate motives. There are different ways in which one can concretely realize multiple zeta values as periods. For example, Goncharov and Manin gave a realization by constructing explicit mixed Tate motives obtained from the moduli spaces $\overline{\mathcal{M}}_{0,n}$ of genus zero curves with marked points [**148**]. Goncharov conjectured that the periods of mixed Tate motives over $\mathbb{Z}$ are exactly the elements of $MZV[1/2\pi i]$, where MZV denotes the $\mathbb{Q}$-vector space $MZV \subset \mathbb{R}$ generated by the multiple zeta values (which is in fact a $\mathbb{Q}$-algebra).
- The result of Belkale and Brosnan [**14**] seems to point to a problem in the expectation that residues of Feynman graphs should be periods of mixed Tate motives. In fact, they show that certain varieties constructed from graphs, on which it is natural to interpret the residues of Feynman integrals as periods, are very general in the sense that they generate the Grothendieck ring of motives. However, more recent results of Bloch, Esnault and Kreimer [**25**] show, for certain classes of graphs, that the Feynman integral, expressed as a period through the Schwinger parameterization, comes from a piece of the cohomology of the graph variety that is actually mixed Tate. This seems to suggest that, although these varieties can give rise to motives that are not mixed Tate, the motivic piece that corresponds to the period relevant for Feynman integrals is still mixed Tate. Recently [**24**], Bloch proposed a lifting at the motivic level of the Connes-Kreimer Hopf algebra.
- Kreimer showed in [**190**] that the period matrix (1.405) of the mixed Hodge–Tate structures associated to the polylogarithms can be seen as

solutions of a Dyson–Schwinger equation, which reveals many of the fundamental structures of renormalization in quantum field theory.

- The technique of renormalization by Birkhoff factorization developed in the Connes–Kreimer theory has been successfully applied to the theory of multiple zeta values in [**158**] and [**212**].

Understanding the precise relation between residues of Feynman integrals and mixed Tate motives remain a question of crucial importance. There are two main obstacles in using the result of Proposition 1.110 and Corollary 1.111 to obtain an explicit identification of the category of flat equisingular vector bundles and a category of mixed Tate motives.

The first problem comes from the fact that, if the homomorphism dual to the homomorphism (1.356) of Theorem 1.106 and Corollary 1.107, mapping $\mathbb{U} \to \mathrm{Difg}(\mathcal{T})$, were coming from a Hopf algebra homomorphism $\mathcal{H}(\mathcal{T}) \to \mathcal{H}_{\mathrm{mot}}$, with $\mathcal{H}_{\mathrm{mot}}$ the Hopf algebra dual to the motive Galois group of mixed Tate motives, then this homomorphism would have to map primitive graphs in $\mathcal{H}(\mathcal{T})$ (i.e. Feynman graphs that are free of subdivergences) to primitive elements in $\mathcal{H}_{\mathrm{mot}}$. However, the results of [**35**] indicate the existence of primitive graphs whose residues are multiple zeta values that do not correspond to primitive elements in $\mathcal{H}_{\mathrm{mot}}$. Also notice that the morphism (1.356) is defined over $\mathbb{C}$ but not over $\mathbb{Q}$, precisely because of the presence of transcendental numbers coming from the residues of graphs.

The second problem lies in the fact that the category of flat equisingular vector bundles is constructed out of irregular-singular connections, while mixed Hodge structures will only arise from regular-singular ones.

This shows that, if Theorem 1.106 and Proposition 1.110 suggest the existence of a direct relation between flat equisingular vector bundles and mixed Tate motives, this relation is more subtle than one might at first imagine. A promising approach appears to be that of approximating irregular-singular connections by regular-singular ones through the phenomenon of "confluence", and then relate the regular-singular differential systems to mixed Hodge structures, hence to mixed Tate motives.

9. The Standard Model of elementary particles

The Standard Model of elementary particle physics is a comprehensive description, based on quantum field theory, of all known elementary particles and their electroweak and strong interactions, in the energy range up to hundreds of GeV (equivalently, a scale of distances down to the order of 10^{-16} centimeters). From the physical point of view, it is an extremely sophisticated theoretical model, which has been tested experimentally to very high precision and accounts better than any other known theory for all known phenomena involving elementary particles. Several attempts have been made at cracking the Standard Model, in the hope of observing new physics beyond what the model accounts for, but at the time of writing (when the new Large Hadron Collider at CERN is still to become operational) no new phenomena in the world of high energy physics beyond the Standard Model have yet been observed, with the important exception of neutrino mixing (cf. §9.6) to which we dedicate special attention in our treatment (cf. §13 below) which will include a model of neutrino mixing with the see-saw mechanism.

Mathematicians may not have much familiarity with the theory of elementary particle physics and with the remarkable conceptual edifice that is the Standard Model.

(There are notable exceptions, see for instance the very interesting book [**188**].)Thus, we attempt to summarize in this section some of the main aspects of the Standard Model. The section can be skipped by the reader who is impatient to jump directly to a more mathematical formulation, but we believe that it is important to give at least a brief overview of some very basic aspects of the physical theory. Our presentation is necessarily brief and somewhat sketchy. For the reader who wishes to learn more about the physics of the Standard Model, there is a wealth of good references available on the subject. Some that we found very useful are [**123**], [**177**], [**267**], [**233**].

Our main purpose is to show that the full Lagrangian of the Standard Model minimally coupled to gravity, in a version that accounts for neutrino mixing, can be *derived* entirely from a very simple mathematical input, using the tools of noncommutative geometry. This will hopefully contribute to providing a clearer conceptual understanding of the wealth of information contained in the Standard Model, in a form which is both palatable to mathematicians and that at the same time can be used to derive specific physical predictions and computations. This was shown in recent work of the authors and Ali Chamseddine [**52**], using techniques of noncommutative geometry previously developed by Connes and Chamseddine–Connes (cf. [**70**], [**45**], [**46**], [**47**], [**48**]). This part of the book is based on [**52**].

The Standard Model comprises two classes of elementary particles: the *fermions* are the basic constituents of matter, while the *bosons* transmit the interactions. Fermions are further subdivided into two classes: *leptons* and *quarks*. The fermions experience electroweak and strong interactions though only the quarks interact through the strong force.
An aspect of the Standard Model that appears somewhat mysterious is the presence of three *generations* of fermions. While the fact that there is an equal number of generations of leptons and quarks can be explained in terms of anomaly cancellation, there is no explanation within the Standard Model that would fix the number of generations.
The Standard Model, in its minimal form, which does not include gravity nor accounts for the phenomenon of neutrino mixing, depends upon 19 parameters. These are not predicted by the model. They are the following:

- 3 charged lepton masses
- 6 quark masses
- 3 gauge coupling constants
- 3 quark mixing angles
- 1 complex phase
- 1 Higgs mass
- 1 coupling constant for the quartic interaction of the Higgs
- 1 QCD vacuum angle

As we discuss in §9.6 below, the observed phenomenon of neutrino mixing is not acccounted for in the original "minimal" standard model and therefore represents the first observed "physics beyond the Standard Model". We discuss the necessary modification to the Standard Model Lagrangian that accounts for the experimental results on neutrino mixing. As we see in §9.6 below, this requires introducing, in addition to the three charged lepton masses, 18 additional real parameters that

correspond to the Dirac and Majorana mass parameters for neutrinos and the angles and phase of the mixing matrix for leptons.
In the case of the Standard Model minimally coupled to gravity, one also has in the Lagrangian the gravitational parameters, which we discuss separately in §9.7 below.

We are going to explain in more detail the role and meaning of all these parameters. In fact, in our model they will all acquire geometric significance and we will be able to derive some relations between them as part of the predictive power of the model.

9.1. Particles and interactions.

Besides electromagnetism, the forces accounted for by the Standard Model are the weak and the strong force. The weak force is responsible for the phenomenon known as neutron β-decay, $n \to pe\bar{\nu}$, whereby a neutron is transformed into a proton, while emitting an electron and an anti-neutrino. This type of phenomenon would not be accounted for by just the theory of electromagnetism. The strong (or nuclear) force is responsible for the fact that atomic nuclei hold together despite the mutual electric repulsion of the protons. It is the elementary constituents of protons and neutrons, namely the quarks, that interact through the strong force. The charge of quarks, with respect to the strong force, is called the color charge. Quarks combine to form composite particles called hadrons (of which the proton and neutron are the two best known examples). There are *gauge bosons* in the Standard Model associated to each of the interactions. The photon is the gauge boson of the electromagnetic force. It is massless and has spin 1. The massive electroweak gauge bosons are the $W^{\pm}$ and Z^0 particles. They also have spin 1. For the strong interaction there are eight gauge bosons, the gluons. They are massless and of spin 1.
The fermions in the Standard Model consist of three generations of leptons and quarks. The first generation has as leptons the electron, which is massive and electrically charged, and the left-handed neutrino ν_e. The quarks in the first generation are the up and down quarks. The second generation consists of the muon and the neutrino ν_μ as leptons and the charmed and strange quarks, and the third generation consists of the leptons τ and the corresponding neutrino ν_τ and of the top and bottom quarks. For each particle there is a corresponding antiparticle. The properties are similar across families: the leptons e, μ, and τ are massive and electrically charged, the neutrinos ν_e, ν_μ and ν_τ are neutral, and are assumed to be massless in the minimal Standard Model. The quarks are massive (with a very broad spectrum of masses) and electrically charged. The quarks also carry a color charge for the strong force.
All electrically charged particles feel the electromagnetic force, all quarks and leptons are subject to the weak interaction, as are the massive gauge bosons of the electroweak force. As we see below, the presence of self-interactions for the gauge bosons depends on the fact that the weak force, unlike electromagnetism, is a nonabelian gauge theory. All quarks and gluons interact through the strong force. Again, the strong force is a nonabelian gauge theory.

Remark 1.113. A comment about notation: in the rest of this chapter we follow standard conventions about notation that are in use in the physics literature, for instance with respect to summation over repeated indices and the raising and lowering of indices. A possible source of confusion is the use of the notation $\bar{\psi}$. In Minkowski space, if a fermion ψ is represented by a column vector, one has $\bar{\psi} = \psi^* \gamma_0$ where γ_0 is a Dirac matrix. Hence this use of the notation $\bar{\psi}$ should not be confused with the ordinary use of the bar notation as complex conjugation, which also occurs frequently in what follows. We hope that the context will clarify the use of the notation.

9.2. Symmetries.

Symmetries play a fundamental role in the Standard Model of elementary particles. There are discrete symmetries, and Lie group symmetries. The discrete symmetries are the CPT symmetries, C for charge, P for parity, and T for time. In terms of the γ matrices, the parity, time reversal, and charge conjugation operators are given on fermions ψ as

$$\begin{aligned} P &: \psi(t,x) \mapsto \gamma^0 \psi(t,-x), \\ T &: \psi(t,x) \mapsto i\gamma^1\gamma^3 \psi(-t,x), \\ C &: \psi(t,x) \mapsto i\gamma^2\gamma^0 \bar{\psi}(t,x). \end{aligned} \tag{1.407}$$

In the case of electromagnetism, the action on the corresponding boson, the photon, is of the form

$$\begin{aligned} P &: A^\mu(t,x) \mapsto A_\mu(t,-x), \\ T &: A^\mu(t,x) \mapsto A_\mu(-t,x), \\ C &: A^\mu(t,x) \mapsto -A^\mu(t,x). \end{aligned} \tag{1.408}$$

The full Standard Model is invariant under the combined CPT transformation. However, unlike electromagnetism and the strong interactions (away from certain special values of quark mixing angles), the electroweak interactions do not preserve the individual C, P and T symmetries. This is related to the fact that, on chiral fermions, parity exchanges right and left chiralities, hence processes which differentiate between different chiralities violate parity.
Among Lie group symmetries there are abelian and nonabelian gauge symmetries. In QED the abelian gauge symmetries corresponding to the group U(1) on fermions and bosons are of the form

$$\psi(t,x) \mapsto \exp(i\lambda(t,x))\,\psi(t,x), \quad A_\mu(t,x) \mapsto A_\mu(t,x) - e^{-1}\partial_\mu \lambda(t,x), \tag{1.409}$$

with e the charge, and $\gamma = \exp(i\lambda)$ a local gauge transformation.
There are nonabelian gauge symmetries by the Lie group SU(3) in the theory of strong interactions. The quarks ψ_j, with color index $j = 1,2,3$, form a 3-dimensional irreducible representation of SU(3), and so do the antiquarks $\bar{\psi}_j$. The gluons A^a_μ, with color index $a = 1,\ldots,8$, form an 8-dimensional irreducible representation of SU(3) which is the adjoint representation. Local gauge transformations U then act by

$$\psi \mapsto U\psi, \quad A_\mu \mapsto U A_\mu U^{-1} + i g^{-1} U \partial_\mu U^{-1}, \tag{1.410}$$

with g the coupling constant.
The U(1) and SU(3) gauge symmetries are full symmetries of the Lagrangian. In particular, this means that there are no mass terms for the U(1) gauge fields and the gluons.

When weak interactions are considered then SU(2) gauge symmetry also plays a role. This is denoted $\mathrm{SU}(2)_L$ and is called the weak isospin symmetry. It is related to phenomena like nuclear β-decay involving charged weak currents. Here the asymmetry between left and right chirality appears. In fact, in the "minimal Standard Model" the left-handed fermions form *weak isodoublets* while the right-handed fermions form *weak isosinglet*. This means that the leptons and quarks are arranged as

$$\begin{aligned} \ell_L &= \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \quad e_R, \\ q_L &= \begin{pmatrix} u \\ d \end{pmatrix}, \quad u_R, \quad d_R. \end{aligned} \tag{1.411}$$

Here we are considering just the minimal case, where there are no right-handed neutrinos. We return later to a detailed discussion of the necessary extensions of the Standard Model that include right-handed neutrinos ν_R, see §9.6, and how they fit in our model, in §13 and following.

Associated to the fermions arranged as in (1.411) as eigenvectors under the representation of $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ there are corresponding *hypercharges*

$$Y_\ell = Y(\ell_L),\ Y_q = Y(q_L),\ Y_u = Y(u_R),\ Y_d = Y(d_R),\ Y_e = Y(e_R). \tag{1.412}$$

These are related to the electric charges Q_{em} and the quantum numbers I_{3L} of the $\mathrm{SU}(2)_L$ representation (i.e. the third generator in the Lie algebra of the weak isospin group) by a linear relation

$$Q_{\mathrm{em}} = \frac{1}{2}Y + I_{3L}. \tag{1.413}$$

There are relations among the hypercharges, which are obtained by imposing anomaly cancellations. These are of the form

$$2Y_q = Y_u + Y_d, \quad Y_q = -\frac{1}{3}Y_\ell, \quad 6Y_q^3 - 3Y_u^3 - 3Y_d^3 + 2Y_\ell^3 - Y_e^3 = 0. \tag{1.414}$$

The values of the hypercharges are given by

$$Y_\ell = -1, \quad Y_e = -2, \quad Y_q = \frac{1}{3}, \quad Y_u = \frac{4}{3}, \quad Y_d = -\frac{2}{3}. \tag{1.415}$$

We discuss, in §13.3 below, how we derive these values for the hypercharges as a consequence of the mathematical model we construct.

A note of warning about conventions. There are other conventions in the physics literature that have the hypercharge equations set to $3Y_q + Y_\ell = 0$, $2Y_q + Y_u + Y_d = 0$ and $6Y_q^3 + 3Y_u^3 + 3Y_d^3 + 2Y_\ell^3 + Y_e^3 = 0$ with hypercharges differing from (1.415) by $Y_e = 2$, $Y_u = -4/3$ and $Y_d = 2/3$. We strictly follow the convention of (1.414) and (1.415).

We return to the discussion of anomalies from a more mathematical perspective later in this chapter; see §19.4, 19.5.

A very important aspect of the theory of the Standard Model is the phenomenon of spontaneous symmetry breaking involving the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetries, which leaves only a resulting $\mathrm{U}(1)_{\mathrm{em}}$ unbroken. The Higgs fields are responsible for the symmetry breaking. In fact, the Higgs potential is of the form

$$V(\mathbf{H}) = -\mu^2|\mathbf{H}|^2 + \lambda|\mathbf{H}|^4. \tag{1.416}$$

The equation for extrema $\mathbf{H}(-\mu^2+2\lambda|\mathbf{H}|^2) = 0$ yields the possible expectation values $\mathbf{H} = 0$ and $|\mathbf{H}|^2 = v^2/2$ with the nonzero order parameter $v = \sqrt{\mu^2/\lambda}$. Imposing this second possibility as a constraint, the original $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ symmetry is reduced to a remaining $\mathrm{U}(1)$ symmetry. The constraint preserves electric charge conservation. The energy scale v is obtained from experiments and governs all the values of the masses.

9.3. Quark mixing: the CKM matrix.

When one accounts for the presence of N generations of matter ($N = 3$ experimentally), one sees that in the part $\mathcal{L}_{Hf}$ of the Standard Model Lagrangian (written in full in §9.4 below, cf. (1.442)) responsible for the coupling of Higgs and fermions, the coupling matrix is not diagonal. That is, the fields are not mass eigenstates. One can transform to mass eigenstates through unitary transformations S_L^u and S_L^d. The expression for the quark charged weak currents then involves the *quark mixing matrix* (or Cabibbo–Kobayashi–Maskawa matrix), which is a product of these unitary matrices $C = S_L^{u*} S_L^d$. In fact, the usual convention in writing the Lagrangian of the Standard Model (cf. §9.4 below) is, instead of using the matrices S_L^u and S_L^d for both up and down quarks, to have the up quarks in diagonal form while the mixing matrix C appears in the down quarks. We are assuming all neutrino masses are zero for the moment, so that we only have the mixing in the quark sector.
For 2 generations, the matrix C depends on just the Cabibbo angle

$$C = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix}, \tag{1.417}$$

while for 3 generations it has the more complicated form

$$C = \begin{pmatrix} C_{ud} & C_{us} & C_{ub} \\ C_{cd} & C_{cs} & C_{cb} \\ C_{td} & C_{ts} & C_{tb} \end{pmatrix}. \tag{1.418}$$

Here the entries depend on 3 angles $\theta_1, \theta_2, \theta_3$ and a complex phase δ, through the expression

$$C = \begin{pmatrix} c_1 & -s_1 c_3 & -s_1 s_3 \\ s_1 c_2 & c_1 c_2 c_3 - s_2 s_3 e_\delta & c_1 c_2 s_3 + s_2 c_3 e_\delta \\ s_1 s_2 & c_1 s_2 c_3 + c_2 s_3 e_\delta & c_1 s_2 s_3 - c_2 c_3 e_\delta \end{pmatrix}, \tag{1.419}$$

for $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, and $e_\delta = \exp(i\delta)$. These parameters can be determined experimentally from weak decays and deep inelastic neutrino scatterings. Recent measurements of the entries of the CKM matrix have been carried out in experiments at Fermilab, in agreement with the predictions of the Standard Model, see [**44**].

9.4. The Standard Model Lagrangian.

In its most explicit form, the Lagrangian density of the *minimal Standard Model* is given by the somewhat uninspiring expression (cf. [**285**]):

$$\mathcal{L}_{SM} = -\tfrac{1}{2}\partial_\nu g^a_\mu \partial_\nu g^a_\mu - g_s f^{abc}\partial_\mu g^a_\nu g^b_\mu g^c_\nu - \tfrac{1}{4} g_s^2 f^{abc} f^{ade} g^b_\mu g^c_\nu g^d_\mu g^e_\nu - \partial_\nu W^+_\mu \partial_\nu W^-_\mu -$$
$$M^2 W^+_\mu W^-_\mu - \tfrac{1}{2}\partial_\nu Z^0_\mu \partial_\nu Z^0_\mu - \tfrac{1}{2c_w^2} M^2 Z^0_\mu Z^0_\mu - \tfrac{1}{2}\partial_\mu A_\nu \partial_\mu A_\nu - igc_w(\partial_\nu Z^0_\mu (W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) -$$
$$Z^0_\nu (W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + Z^0_\mu (W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) - igs_w(\partial_\nu A_\mu (W^+_\mu W^-_\nu -$$
$$W^+_\nu W^-_\mu) - A_\nu (W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + A_\mu (W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) -$$
$$\tfrac{1}{2} g^2 W^+_\mu W^-_\mu W^+_\nu W^-_\nu + \tfrac{1}{2} g^2 W^+_\mu W^-_\nu W^+_\mu W^-_\nu + g^2 c_w^2 (Z^0_\mu W^+_\mu Z^0_\nu W^-_\nu - Z^0_\mu Z^0_\mu W^+_\nu W^-_\nu) +$$

$$
\begin{gathered}
g^2 s_w^2 (A_\mu W_\mu^+ A_\nu W_\nu^- - A_\mu A_\mu W_\nu^+ W_\nu^-) + g^2 s_w c_w (A_\mu Z_\nu^0 (W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-) - \\
2 A_\mu Z_\mu^0 W_\nu^+ W_\nu^-) - \tfrac{1}{2} \partial_\mu H \partial_\mu H - \tfrac{1}{2} m_h^2 H^2 - \partial_\mu \phi^+ \partial_\mu \phi^- - M^2 \phi^+ \phi^- - \tfrac{1}{2} \partial_\mu \phi^0 \partial_\mu \phi^0 - \tfrac{1}{2c_w^2} M^2 \phi^0 \phi^0 - \\
\beta_h \left(\tfrac{2M^2}{g^2} + \tfrac{2M}{g} H + \tfrac{1}{2} (H^2 + \phi^0 \phi^0 + 2\phi^+ \phi^-) \right) + \tfrac{2M^4}{g^2} \alpha_h - g \alpha_h M \left(H^3 + H \phi^0 \phi^0 + 2H \phi^+ \phi^- \right) - \\
\tfrac{1}{8} g^2 \alpha_h \left(H^4 + (\phi^0)^4 + 4(\phi^+ \phi^-)^2 + 4(\phi^0)^2 \phi^+ \phi^- + 4H^2 \phi^+ \phi^- + 2(\phi^0)^2 H^2 \right) - g M W_\mu^+ W_\mu^- H - \\
\tfrac{1}{2} g \tfrac{M}{c_w^2} Z_\mu^0 Z_\mu^0 H - \tfrac{1}{2} i g \left(W_\mu^+ (\phi^0 \partial_\mu \phi^- - \phi^- \partial_\mu \phi^0) - W_\mu^- (\phi^0 \partial_\mu \phi^+ - \phi^+ \partial_\mu \phi^0) \right) + \\
\tfrac{1}{2} g \left(W_\mu^+ (H \partial_\mu \phi^- - \phi^- \partial_\mu H) + W_\mu^- (H \partial_\mu \phi^+ - \phi^+ \partial_\mu H) \right) + \tfrac{1}{2} g \tfrac{1}{c_w} Z_\mu^0 (H \partial_\mu \phi^0 - \phi^0 \partial_\mu H) - \\
i g \tfrac{s_w^2}{c_w} M Z_\mu^0 (W_\mu^+ \phi^- - W_\mu^- \phi^+) + i g s_w M A_\mu (W_\mu^+ \phi^- - W_\mu^- \phi^+) - i g \tfrac{1 - 2c_w^2}{2c_w} Z_\mu^0 (\phi^+ \partial_\mu \phi^- - \\
\phi^- \partial_\mu \phi^+) + i g s_w A_\mu (\phi^+ \partial_\mu \phi^- - \phi^- \partial_\mu \phi^+) - \tfrac{1}{4} g^2 W_\mu^+ W_\mu^- \left(H^2 + (\phi^0)^2 + 2\phi^+ \phi^- \right) - \\
\tfrac{1}{8} g^2 \tfrac{1}{c_w^2} Z_\mu^0 Z_\mu^0 \left(H^2 + (\phi^0)^2 + 2(2s_w^2 - 1)^2 \phi^+ \phi^- \right) - \tfrac{1}{2} g^2 \tfrac{s_w^2}{c_w} Z_\mu^0 \phi^0 (W_\mu^+ \phi^- + W_\mu^- \phi^+) - \\
\tfrac{1}{2} i g^2 \tfrac{s_w^2}{c_w} Z_\mu^0 H (W_\mu^+ \phi^- - W_\mu^- \phi^+) + \tfrac{1}{2} g^2 s_w A_\mu \phi^0 (W_\mu^+ \phi^- + W_\mu^- \phi^+) + \tfrac{1}{2} i g^2 s_w A_\mu H (W_\mu^+ \phi^- - \\
W_\mu^- \phi^+) - g^2 \tfrac{s_w}{c_w} (2c_w^2 - 1) Z_\mu^0 A_\mu \phi^+ \phi^- - g^2 s_w^2 A_\mu A_\mu \phi^+ \phi^- + \tfrac{1}{2} i g_s \lambda_{ij}^a (\bar{q}_i^\sigma \gamma^\mu q_j^\sigma) g_\mu^a - \bar{e}^\lambda (\gamma \partial + m_e^\lambda) e^\lambda - \\
\bar{\nu}^\lambda \gamma \partial \nu^\lambda - \bar{u}_j^\lambda (\gamma \partial + m_u^\lambda) u_j^\lambda - \bar{d}_j^\lambda (\gamma \partial + m_d^\lambda) d_j^\lambda + i g s_w A_\mu \left(-(\bar{e}^\lambda \gamma^\mu e^\lambda) + \tfrac{2}{3} (\bar{u}_j^\lambda \gamma^\mu u_j^\lambda) - \tfrac{1}{3} (\bar{d}_j^\lambda \gamma^\mu d_j^\lambda) \right) + \\
\tfrac{ig}{4c_w} Z_\mu^0 \{ (\bar{\nu}^\lambda \gamma^\mu (1 + \gamma^5) \nu^\lambda) + (\bar{e}^\lambda \gamma^\mu (4s_w^2 - 1 - \gamma^5) e^\lambda) + (\bar{d}_j^\lambda \gamma^\mu (\tfrac{4}{3} s_w^2 - 1 - \gamma^5) d_j^\lambda) + (\bar{u}_j^\lambda \gamma^\mu (1 - \\
\tfrac{8}{3} s_w^2 + \gamma^5) u_j^\lambda) \} + \tfrac{ig}{2\sqrt{2}} W_\mu^+ \left((\bar{\nu}^\lambda \gamma^\mu (1 + \gamma^5) e^\lambda) + (\bar{u}_j^\lambda \gamma^\mu (1 + \gamma^5) C_{\lambda\kappa} d_j^\kappa) \right) + \\
\tfrac{ig}{2\sqrt{2}} W_\mu^- \left((\bar{e}^\lambda \gamma^\mu (1 + \gamma^5) \nu^\lambda) + (\bar{d}_j^\kappa C_{\kappa\lambda}^\dagger \gamma^\mu (1 + \gamma^5) u_j^\lambda) \right) + \\
\tfrac{ig}{2\sqrt{2}} \tfrac{m_e^\lambda}{M} \left(-\phi^+ (\bar{\nu}^\lambda (1 - \gamma^5) e^\lambda) + \phi^- (\bar{e}^\lambda (1 + \gamma^5) \nu^\lambda) \right) - \tfrac{g}{2} \tfrac{m_e^\lambda}{M} \left(H (\bar{e}^\lambda e^\lambda) + i \phi^0 (\bar{e}^\lambda \gamma^5 e^\lambda) \right) + \\
\tfrac{ig}{2M\sqrt{2}} \phi^+ \left(-m_d^\kappa (\bar{u}_j^\lambda C_{\lambda\kappa} (1 - \gamma^5) d_j^\kappa) + m_u^\lambda (\bar{u}_j^\lambda C_{\lambda\kappa} (1 + \gamma^5) d_j^\kappa) \right) + \\
\tfrac{ig}{2M\sqrt{2}} \phi^- \left(m_d^\lambda (\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger (1 + \gamma^5) u_j^\kappa) - m_u^\kappa (\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger (1 - \gamma^5) u_j^\kappa) \right) - \tfrac{g}{2} \tfrac{m_u^\lambda}{M} H (\bar{u}_j^\lambda u_j^\lambda) - \\
\tfrac{g}{2} \tfrac{m_d^\lambda}{M} H (\bar{d}_j^\lambda d_j^\lambda) + \tfrac{ig}{2} \tfrac{m_u^\lambda}{M} \phi^0 (\bar{u}_j^\lambda \gamma^5 u_j^\lambda) - \tfrac{ig}{2} \tfrac{m_d^\lambda}{M} \phi^0 (\bar{d}_j^\lambda \gamma^5 d_j^\lambda) + \bar{G}^a \partial^2 G^a + g_s f^{abc} \partial_\mu \bar{G}^a G^b g_\mu^c + \\
\bar{X}^+ (\partial^2 - M^2) X^+ + \bar{X}^- (\partial^2 - M^2) X^- + \bar{X}^0 (\partial^2 - \tfrac{M^2}{c_w^2}) X^0 + \bar{Y} \partial^2 Y + i g c_w W_\mu^+ (\partial_\mu \bar{X}^0 X^- - \\
\partial_\mu \bar{X}^+ X^0) + i g s_w W_\mu^+ (\partial_\mu \bar{Y} X^- - \partial_\mu \bar{X}^+ Y) + i g c_w W_\mu^- (\partial_\mu \bar{X}^- X^0 - \partial_\mu \bar{X}^0 X^+) + \\
i g s_w W_\mu^- (\partial_\mu \bar{X}^- Y - \partial_\mu \bar{Y} X^+) + i g c_w Z_\mu^0 (\partial_\mu \bar{X}^+ X^+ - \partial_\mu \bar{X}^- X^-) + i g s_w A_\mu (\partial_\mu \bar{X}^+ X^+ - \\
\partial_\mu \bar{X}^- X^-) - \tfrac{1}{2} g M \left(\bar{X}^+ X^+ H + \bar{X}^- X^- H + \tfrac{1}{c_w^2} \bar{X}^0 X^0 H \right) + \\
\tfrac{1 - 2c_w^2}{2c_w} i g M \left(\bar{X}^+ X^0 \phi^+ - \bar{X}^- X^0 \phi^- \right) + \tfrac{1}{2c_w} i g M \left(\bar{X}^0 X^- \phi^+ - \bar{X}^0 X^+ \phi^- \right) + \\
i g M s_w \left(\bar{X}^0 X^- \phi^+ - \bar{X}^0 X^+ \phi^- \right) + \tfrac{1}{2} i g M \left(\bar{X}^+ X^+ \phi^0 - \bar{X}^- X^- \phi^0 \right) .
\end{gathered}
$$

Here the notation is as follows.

- Gauge bosons: A_μ, $W_\mu^\pm$, Z_μ^0, g_μ^a
- Quarks: u_j^κ, d_j^κ, collectively denoted q_j^κ.
- Leptons: e^κ, ν^κ
- Higgs fields: H, ϕ^0, ϕ^+, ϕ^-
- Masses: m_d^κ, m_u^κ, m_e^κ, m_h (Higgs mass) and M (mass of the W)
- c_w and s_w: cosine and sine of the weak mixing angle θ_w
- $C_{\lambda\kappa}$: Cabibbo–Kobayashi–Maskawa mixing matrix
- Coupling constants: g (with $\alpha = s_w^2 g^2/(4\pi)$ the fine structure constant) g_s (coupling of the strong force), $\alpha_h = \frac{m_h^2}{4M^2}$ (Higgs scattering parameter)
- β_h: tadpole constant
- f^{abc}: structure constants of the Lie algebra of SU(3)
- λ_{ij}^a: Gell-Mann matrices (1.421)
- Ghosts: G^a, X^0, X^+, X^-, Y
- The choice of gauge fixing is the Feynman gauge for all gauge fields except the SU(2) ones and the Feynman–'t Hooft gauge (cf. §9.5 below) for the SU(2) gauge fields.

For quarks q_j^κ, the index κ is the generation (flavor) index. Thus, for example, when we write u^κ with $\kappa = 2$, this stands for the charmed quark. The index j is the color index.
In the case of leptons, again κ is the generation index, so that, for example e^κ for $\kappa = 2$ is the muon and ν^κ the corresponding neutrino. In the minimal Standard Model one assumes that the neutrinos ν^κ only appear in their left-handed chirality ν_L^κ and are massless. This hypothesis will be modified in the actual model we work with.
The coupling constants are expressed also in the form $g_1 = g\tan(\theta_w)$ (electromagnetic coupling), $g_2 = g$ (weak coupling) and $g_3 = g_s$ (strong coupling). One correspondingly defines

$$\alpha_i = \frac{g_i^2}{4\pi}, \quad \text{where} \quad \alpha_{\text{em}} = s_w^2 \alpha_2 \tag{1.420}$$

is the fine structure constant.
The structure constants f^{abc}, with the indices $a, b, c = 1, \ldots, 8$, of the Lie algebra of SU(3) satisfy the relations

$$[g_a, g_b] = i f^{abc} g_c,$$

where the g_a are the generators of the Lie algebra. The f^{abc} are completely antisymmetric in the three indices and are specified by the values

$$f^{123} = 2, \quad f^{147} = f^{165} = f^{246} = f^{257} = f^{376} = 1, \quad f^{458} = f^{678} = \sqrt{3}.$$

The Gell-Mann matrices give the fundamental representation of the Lie algebra of SU(3) in traceless 3×3 matrices. They are defined as follows.

DEFINITION 1.114. *The Gell-Mann matrices λ_{ij}^a are given by*

$$\begin{aligned}
\lambda_{ij}^1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_{ij}^2 = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_{ij}^3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_{ij}^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\
\lambda_{ij}^5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_{ij}^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
\lambda_{ij}^7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_{ij}^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{1.421}$$

The matrices (1.421) are self-adjoint and satisfy the relation

$$\operatorname{Tr}(\lambda^a \lambda^b) = 2\delta^{ab}, \tag{1.422}$$

with δ^{ab} the Kronecker delta.

Here we have chosen to write the Standard Model Lagrangian exactly as given in Veltman's book [**285**], including the somewhat unusual choice of gauge fixing, with the Feynman gauge replaced by the Feynman–'t Hooft gauge for the SU(2) gauge fields (see the discussion in §9.5). The reason for maintaining the same notation and conventions of [**285**] is that this is a standard reference in the physics literature for the full Lagrangian of the Standard Model. In the course of developing our mathematical model in the coming sections, we will also translate the notations from the physical ones to those of noncommutative geometry.

One of our goals in this chapter is to arrive at a conceptually clean mathematical understanding of these data, which makes it possible to recover the full complexity of the expression above starting from very simple principles and as small as possible an input of data. This will be achieved through the *spectral action principle* for a noncommutative spectral geometry $(\mathcal{A}, \mathcal{H}, \mathcal{D})$ (cf. §11). In fact, we not only recover the original Standard Model, but we can directly deal with the refined form that incorporates neutrino masses and mixing as well as the see-saw phenomenon (cf. [**233**] §7.2.4).

For the moment, we just introduce some notation, in order to write the expression above in a more concise way. The terms of $\mathcal{L}_{SM}$ can be grouped together in six different types, so that we can write the Lagrangian more synthetically as

$$\mathcal{L}_{SM} = \mathcal{L}_g + \mathcal{L}_{gH} + \mathcal{L}_H + \mathcal{L}_{gf} + \mathcal{L}_{Hf} + \mathcal{L}_{ghosts}. \tag{1.423}$$

The term $\mathcal{L}_g$ denotes the pure gauge boson part, where we write g for the spin-1 bosons, that is, the eight gluons g^a_μ, the photon A_μ, and the intermediate bosons $W^\pm_\mu$ and Z^0_μ. The term $\mathcal{L}_{gH}$ gives the minimal coupling of the spin-1 bosons with the Higgs fields, namely the spin-0 bosons H, ϕ^0, ϕ^+, ϕ^-. The term $\mathcal{L}_H$ contains the quartic Higgs self-interaction. In addition to the coupling constants for the gauge fields, the fermion kinetic term $\mathcal{L}_{gf}$ contains the coupling of fermions with the gauge bosons which involves the hypercharges Y_L, Y_R. The term $\mathcal{L}_{Hf}$ contains the Yukawa coupling of the Higgs fields with the fermions. In these two terms f stands for the spin-1/2 fermions, namely all quarks u^κ_j, d^κ_j, with the upper generation index and the lower color index, and leptons e^κ, ν^κ with the upper generation index. We then have the remaning term $\mathcal{L}_{ghosts}$, which contains the kinetic terms of the Faddeev–Popov ghosts G^a, X^0, X^+, X^-, and Y, as well as their interactions with the gauge and Higgs fields. All these ghost terms are gathered at the end of the long explicit formula for $\mathcal{L}_{SM}$. They have the important property that they can be obtained canonically from the other five terms $\mathcal{L}_g + \mathcal{L}_{gH} + \mathcal{L}_H + \mathcal{L}_{gf} + \mathcal{L}_{Hf}$, using the general technique of gauge fixing, which we explain briefly in §9.5 below. The difference between the Feynman and Feynman–'t Hooft gauge in the gauge fixing is also discussed in §9.5 below. The gauge fixing term $\mathcal{L}_{\text{fix}}$ is then combined with the kinetic terms for the gauge fields.

9.5. Quantum level: anomalies, ghosts, gauge fixing.

Up to now we have only introduced the Standard Model at the classical level. However, to make computations that can be experimentally tested one needs not only the tree level approximation but one needs to incorporate the quantum corrections coming from the higher order terms, namely from Feynman diagrams with loops. It is computationally very challenging to perform such calculations, even for a small number of loops. It is in this context that the general tools of perturbative quantum field theory and the theory of renormalization discussed earlier in this chapter become very significant.

The Feynman rules for QED assign the fermion and photon propagators, and the contribution of one type of interaction vertex, which corresponds to the triangle vertex of the fermion photon interaction. The Feynman rules for QCD assign the quark, gluon, and ghost propagators and the contributions of various types of interaction vertices: the triangle vertices corresponding to a three-gluon vertex, a quark-gluon vertex, or a ghost-gluon vertex, and a valence four four-gluon vertex. In the electroweak case, the Feynman rules are still more complicated and they involve more propagators and interaction vertices: the propagators associated to the Z and W bosons, to the charged Higgs, vertices of fermion-boson interaction, etc (see e.g. [**167**]). In our discussion we will only encounter explicit quantum correction calculations when we use the renormalization group equation and the corresponding beta functions at one or two loops to obtain predictions at ordinary energies from the relations that can be derived in our model at unification energies, see §17.2, §17.6, §17.10.

An interesting phenomenon that appears when one considers higher order corrections is that certain apparent symmetries of the Lagrangian turn out not to be symmetries of the theory. This is due to the presence of *anomalies*. The typical case, which we return to discuss from the point of view of noncommutative geometry later in this chapter, is the *axial* U(1) *symmetry*. The QCD Lagrangian (if quarks were massless) would have a symmetry

$$\psi \mapsto \exp(-i\theta\gamma_5)\,\psi. \tag{1.424}$$

However, this symmetry does not survive at the quantum level due to anomalies. Anomalies can be thought of as a problem with change of variable in path integrals, by the effect of which what is classically a conserved current (by Noether's theorem) associated to the symmetry is no longer divergence-free due to quantum corrections. The divergence free condition for the classical Noether current is reformulated in terms of Ward identities between Green's functions.

In general the Ward identities determine relations between the Green's functions arising from symmetries of the Lagrangian. These influence the renormalizability of theories with nontrivial symmetries, by ensuring cancellations of divergences between different sectors of the theory. There is a very interesting and subtle interplay between symmetries, Ward identities, and perturbative renormalization, see e.g. [**60**]. We discussed in the previous part of this chapter a mathematical formulation of perturbative renormalization based on Hopf algebras, Birkhoff factorization, equisingular connections and affine group schemes. The compatibility of the Connes–Kreimer theory and this general formalism with gauge transformations and Ward identities was recently proved by van Suijlekom for QED in [**276**] and in the more general case of nonabelian gauge theories in [**277**]. A subtle point is that the Ward identities can be affected by higher order corrections, which can generate anomalous terms. In gauge theories, the problem arises primarily from fermions coupling to gauge fields, as in the Adler–Bell–Jackiw (ABJ) anomaly which manifests itself in the renormalization of 1-loop diagrams (triangle fermion loops).

In the DimReg regularization scheme, the ABJ anomaly is closely related to the problem of defining γ_5 for complex dimensions (cf. [**176**]). We discuss later how this can be formulated in the context of noncommutative geometry, in §19 below.

In the case of electroweak interactions where the vector boson associated to the weak interaction acquires mass through the Higgs mechanism, renormalizability is achieved by showing that the theory is obtained via a formal transformation of the fields from a renormalizable gauge theory (the latter can have non-physical particles that may spoil unitarity, cf. [**168**]). For such a formal transformation to be possible, one needs to be able to renormalize the gauge theory in a gauge invariant way. This can be spoiled by the presence of the ABJ anomaly, as shown in [**152**]. Namely, in the case of interactions involving the Noether current of the classical axial symmetry (1.424), at the level of 1-loop quantum corrections, triangular diagrams do not preserve the current conservation that exists at the tree level. Unless the anomalies cancel out the theory ceases to be renormalizable. The cancellation of anomalies gives an important constraint on the theory. The question of constructing gauge theories that are free of anomalies, both for electroweak and strong interactions, was considered in [**141**], where conditions are given on which gauge groups and representations will satisfy the vanishing of the ABJ anomaly for triangle diagrams.

In the case of gauge theories, the Ward identities (or Slavnov–Taylor identities) are related to the symmetries by Becchi-Rouet-Stora (BRS) gauge transformations. For instance, in the SU(2) case, the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F^a_{\mu\nu} F^{a\,\mu\nu} + \bar{\psi} i \gamma^\mu D_\mu \psi - m \bar{\psi}\psi, \tag{1.425}$$

with $D_\mu \psi = (\partial_\mu - igA^a_\mu T^a)\psi$ and $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\epsilon^{abc} A^b_\mu A^c_\nu$, is invariant under local gauge transformations

$$\delta\psi = -iT^a \theta^a \psi \quad \text{and} \quad \delta A^a_\mu = \epsilon^{abc}\theta^b A^c_\mu - g^{-1}\partial_\mu \theta^a. \tag{1.426}$$

One considers the additional gauge fixing term and Faddeev–Popov ghost terms, which give the resulting Lagrangian

$$\widetilde{\mathcal{L}} = \mathcal{L} + \mathcal{L}_{\rm fix} + \mathcal{L}_{ghost}, \tag{1.427}$$

where

$$\begin{aligned} \mathcal{L}_{\rm fix} = &\ \tfrac{-1}{2\xi}(\partial_\mu A^a_\mu)^2 \\ \mathcal{L}_{ghost} = &\ ic^*_a \partial^\mu (\delta_{ab}\partial_\mu - g\epsilon_{abc}A^c_\mu) c_b, \end{aligned} \tag{1.428}$$

with c the ghost field.

In the case described above of the Standard Model Lagrangian, the gauge is chosen to be the *Feynman gauge*, which is as in (1.428) with $\xi = 1$, for all the gauge fields except for the SU(2) gauge fields. One advantage of the choice $\xi = 1$ is that the kinetic terms for the gauge fields in (1.425) are simply given by

$$-\frac{1}{2}\partial_\nu A^a_\mu \, \partial_\nu A^a_\mu \tag{1.429}$$

The gauge chosen for the SU(2) gauge field W^μ_a in the Standard Model is the *Feynman–'t Hooft gauge*, of the form

$$\mathcal{L}_{\rm fix} = -\frac{1}{2}\mathcal{C}^2, \quad \text{with} \quad \mathcal{C}_a = -\partial_\mu W^\mu_a + M_a \, \phi_a. \tag{1.430}$$

Here M_a is the mass of the massive vector boson W_a. The change of variables from the isospin indices W_a, ϕ_a to the standard $(Z, W^\pm)$, $(\phi^0, \phi^\pm)$ notation is given by

$$\begin{aligned} Z^0 = &\ W_0\,, \quad W^+ = \tfrac{W_1 - iW_2}{\sqrt{2}}\,, \quad W^- = \tfrac{W_1 + iW_2}{\sqrt{2}}\,, \\ \phi^0 = &\ \phi_0\,, \quad \phi^+ = \tfrac{\phi_1 - i\phi_2}{\sqrt{2}}\,, \quad \phi^- = \tfrac{\phi_1 + i\phi_2}{\sqrt{2}}\,. \end{aligned} \tag{1.431}$$

Thus, one can expand the gauge fixing term in the form

$$\begin{aligned} \mathcal{L}_{\text{fix}} = & -\tfrac{1}{2}(\partial_\mu W_a^\mu)^2 - \tfrac{1}{2c_w^2} M^2 \phi^0 \phi^0 - M^2 \phi^+ \phi^- \\ & + M(\tfrac{1}{c_w}\phi^0 \partial_\mu Z_\mu^0 + \phi^- \partial_\mu W_\mu^+ + \phi^+ \partial_\mu W_\mu^-). \end{aligned} \tag{1.432}$$

We separate the terms of this expansion as

$$\mathcal{L}_{\text{fix}} = \mathcal{L}_{feyn} + \mathcal{L}_{\text{fix}}^M + \mathcal{L}_{cross}, \tag{1.433}$$

where the cross terms form the last group.

The ghost terms were gathered at the end of the formula for the full Lagrangian in the form

$$\begin{aligned} \mathcal{L}_{ghost} = & \bar{G}^a \partial^2 G^a + g_s f^{abc} \partial_\mu \bar{G}^a G^b g_\mu^c \\ & + \bar{X}^+(\partial^2 - M^2) X^+ + \bar{X}^-(\partial^2 - M^2) X^- \\ & + \bar{X}^0(\partial^2 - \tfrac{M^2}{c_w^2}) X^0 + \bar{Y} \partial^2 Y \\ & + igc_w W_\mu^+(\partial_\mu \bar{X}^0 X^- - \partial_\mu \bar{X}^+ X^0) \\ & + igs_w W_\mu^+(\partial_\mu \bar{Y} X^- - \partial_\mu \bar{X}^+ Y) \\ & + igc_w W_\mu^-(\partial_\mu \bar{X}^- X^0 - \partial_\mu \bar{X}^0 X^+) \\ & + igs_w W_\mu^-(\partial_\mu \bar{X}^- Y - \partial_\mu \bar{Y} X^+) \\ & + igc_w Z_\mu^0(\partial_\mu \bar{X}^+ X^+ - \partial_\mu \bar{X}^- X^-) \\ & + igs_w A_\mu(\partial_\mu \bar{X}^+ X^+ - \partial_\mu \bar{X}^- X^-) \\ & - \tfrac{1}{2} gM \left(\bar{X}^+ X^+ H + \bar{X}^- X^- H + \tfrac{1}{c_w^2} \bar{X}^0 X^0 H \right) \\ & + \tfrac{1-2c_w^2}{2c_w} igM \left(\bar{X}^+ X^0 \phi^+ - \bar{X}^- X^0 \phi^- \right) \\ & + \tfrac{1}{2c_w} igM \left(\bar{X}^0 X^- \phi^+ - \bar{X}^0 X^+ \phi^- \right) \\ & + igMs_w \left(\bar{X}^0 X^- \phi^+ - \bar{X}^0 X^+ \phi^- \right) \\ & + \tfrac{1}{2} igM \left(\bar{X}^+ X^+ \phi^0 - \bar{X}^- X^- \phi^0 \right). \end{aligned} \tag{1.434}$$

We ignore them in the analysis of the Lagrangian, since they are obtained in a systematic manner from the other terms. The appearance of the Feynman gauge fixing terms $\mathcal{L}_{\text{fix}}$ will hardly be noticeable, since in the Feynman gauge they are almost invisible, by just changing the summation on indices in the gauge kinetic terms. On the contrary, the appearance of the Feynman–'t Hooft gauge fixing terms will be quite noticeable. In fact, after an integration by parts, this gauge fixing amounts to replacing

$$M\left(\frac{1}{c_w} Z_\mu^0 \partial_\mu \phi^0 + W_\mu^+ \partial_\mu \phi^- + W_\mu^- \partial_\mu \phi^+\right) \quad \mapsto \quad -\frac{1}{2c_w^2} M^2 \phi^0 \phi^0 - M^2 \phi^+ \phi^-, \tag{1.435}$$

i.e. replacing a minimal coupling term by a mass term for the Goldstone boson components of the Higgs field.

The resulting action functional is no longer gauge invariant, but is invariant under the BRS generalized gauge transformations. These are transformations of the form (1.426), where θ^a is a specific type of gauge function of the form

$$\theta^a = -g\omega\sigma^a, \tag{1.436}$$

where ω is a spacetime-independent anticommuting Graßmann variable and σ^a is the imaginary component of the ghost field, written in the form $c_a = (\rho_a + i\sigma_a)/\sqrt{2}$. The BRS transformations also act on the ghost field as

$$\delta\rho^a = -i\omega\partial^\mu A_\mu^a/\xi \quad \text{and} \quad \delta\sigma^a = -g\omega\epsilon^{abc}\sigma^b\sigma^c/2. \tag{1.437}$$

9.6. Massive neutrinos.

A first step beyond the physics of the Standard Model is needed in order to account for neutrino masses. Experimental evidence for neutrino masses comes from neutrino oscillations. The phenomenon is observed in atmospheric neutrinos produced by the interaction of cosmic ray protons with atomic nuclei in the upper atmosphere. These interactions produce electron and muon neutrinos. The experimental count of the ratio of these two types of neutrinos however is not in agreement with the expected ratio. This is understood as a sign of the presence of oscillations between the different flavors, i.e. the fact the presence of mass eigenstates that are a superposition of muon and electron types. In fact, if neutrinos are no longer assumed to be massless, then, just as in the case of quark mixing, the mass eigenstates will be a mixture of the three generations ν_e, ν_μ, ν_τ. This is *neutrino mixing* and, as in the case of the CKM matrix for quarks, it is governed by a neutrino mixing matrix, the PMNS (Pontecorvo–Maki–Nakagawa–Sakata) matrix. Neutrino oscillations are seen as the experimental signature of the presence of neutrino mixing. Another evidence for neutrino oscillations comes from the so-called "solar neutrino puzzle", by which the amount of solar neutrinos detected on Earth appeared to be in defect of the theoretical models for solar emission, until it was found that a fraction of the electron neutrinos emitted by the Sun seems to convert into other flavors before reaching the Earth. For a detailed discussion of experimental evidence of neutrino masses and its theoretical implications, see [**233**].
In the "minimal Standard Model", as we have seen above, neutrinos are massless. The observations of neutrino oscillations therefore represent the first source of evidence of physics beyond the Standard Model. There are various models that extend the minimal Standard Model we discussed above and account for neutrino mixing. Most require expanding the field content of the minimal Standard Model. This is the case, for instance, in the so called Left-Right symmetric models, where the group $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ of the electroweak theory is replaced by the larger $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R \times \mathrm{U}(1)_{B-L}$ and the relation (1.413) is replaced by

$$Q = I_{3L} + I_{3R} + \frac{B-L}{2}. \tag{1.438}$$

The Higgs sector also changes in general in such models (cf. [**233**], §7 and §8). The model that we obtain from the mathematical input we describe in section §13 below is in fact the minimal extension of the Standard Model that incorporates neutrino mixing and has a see-saw mechanism (see §17.5 below) that accounts for the fact that the observed bounds on neutrino masses makes them so much lighter than the charged fermions.
In this type of model neutrino masses are incorporated by changing the fermionic content of the minimal Standard Model by adding right-handed neutrinos. This brings about a necessary discussion on the description of neutrinos as Dirac or Majorana fermions.

9.6.1. *Dirac and Majorana fermions.*

For the moment we assume that we work in Minkowski space. Suppose given a fermion ψ_L, of left-handed chirality, i.e. such that $\frac{1}{2}(1-\gamma_5)\psi_L = \psi_L$, where $\frac{1}{2}(1-\gamma_5)$ and $\frac{1}{2}(1+\gamma_5)$ are the chirality projections. Let $\hat{\psi}_R$ denote the image of ψ_L under the composite CPT symmetry of (1.407), that is, the right-handed state of the antiparticle. Similarly, one has states ψ_R and $\hat{\psi}_L$. If the fermion is charged, then $\hat{\psi}_R$ and ψ_R have opposite charge and same helicity. In particular, if the fermion ψ is a massive particle, hence it moves at speeds strictly smaller than the speed of light, a boosted observer that overtakes the velocity of a particle ψ_L will then see a right-handed object, but with the same charge. In the absence of charge, however, one has the possibility of identifying the states $\hat{\psi}_R$ and ψ_R and similarly the states ψ_L and $\hat{\psi}_L$. Spinors for which this identification holds are called *Majorana spinors*. If one does not have this identification, one talks about *Dirac spinors*. These have twice as many degrees of freedom as Majorana spinors.

In the minimal Standard Model, the left-handed neutrinos ν_L are relativistic particles, that is, being massless they move at light speed. In this case, it makes no sense to talk about the possibility of a different chirality seen by a boosted observer. One says that the ν_L of the minimal Standard Model are Weyl spinors.

In extensions of the Standard Model where neutrinos have masses, one needs to consider also the right-handed state ν_R seen by a boosted observer and similarly one has the $\hat{\nu}_L$ and $\hat{\nu}_R$ states. Since neutrinos do not have charge, it is possible to either consider these as different states, or identify $\nu_L = \hat{\nu}_L$ and $\nu_R = \hat{\nu}_R$, so that neutrinos are their own antiparticles. It is this description of neutrinos as Majorana spinors that will appear naturally in our model.

Our model will be derived after a Wick rotation to Euclidean signature. It is well known that, unlike Dirac spinors, Majorana spinors do not have a Euclidean version. This seems to pose a serious problem with the picture we just described for massive neutrinos. However, as we discuss in detail in §16.2 below, we can use a formalism based on the Pfaffian and Graßmann variables to provide a substitute for the formalism of Majorana spinors in the Euclidean setup. This role of the Pfaffian is quite well known in the physics literature; see for instance [**171**], [**184**], [**237**].

For Majorana spinors, the mass terms in the Lagrangian can be written in the form (cf. [**233**], §4.5)

$$\frac{1}{2}\sum_{\lambda\kappa} \bar{\psi}_{\lambda L} \mathcal{S}_{\lambda\kappa} \hat{\psi}_{\kappa R} + \frac{1}{2}\sum_{\lambda\kappa} \overline{\bar{\psi}_{\lambda L} \mathcal{S}_{\lambda\kappa} \hat{\psi}_{\kappa R}}. \tag{1.439}$$

Here $\mathcal{S}_{\kappa\lambda}$ is a symmetric matrix. In particular, in the special case where this matrix is of the form

$$\mathcal{S}_{\kappa\lambda} = \begin{pmatrix} 0 & T \\ T & 0 \end{pmatrix},$$

one can identify a pair of Majorana spinors with a Dirac spinor and obtain a Dirac mass term. In the general form, where $\mathcal{S}_{\lambda\kappa}$ also has a diagonal part, we will see that the relative size of the terms in the matrix $\mathcal{S}$ will be the source of the see-saw mechanism (see §17.5 below).

One sometimes refers to the off-diagonal part of $\mathcal{S}_{\lambda\kappa}$ as the Dirac mass terms and to the diagonal part as the Majorana mass terms. For a pair ψ and χ of Majorana fermions, the Dirac mass terms correspond to cross terms of the form $T\bar{\psi}_L\hat{\chi}_R$ and $T\bar{\chi}_L\hat{\psi}_R$ and Hermitian conjugate, while the Majorana mass terms correspond to cross terms involving $\bar{\psi}_L\hat{\psi}_R$ and $\bar{\chi}_L\hat{\chi}_R$. We see in the next subsection how such terms appear in the modified Standard Model Lagrangian.

9.6.2. *The modified Standard Model Lagrangian.*

We now describe more precisely the modification to the Standard Model Lagrangian of §9.4 that will be relevant to our setting and which incorporates massive neutrinos, the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) mixing matrix, and the Majorana mass term for neutrinos.
The terms of the original Lagrangian of §9.4 that need to be modified are those of the form

$$\begin{array}{c} -\bar{\nu}^\lambda \gamma\partial\nu^\lambda \\ +\frac{ig}{2\sqrt{2}} W^+_\mu \left(\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5)e^\lambda\right) \\ +\frac{ig}{2\sqrt{2}} W^-_\mu \left(\bar{e}^\lambda\gamma^\mu(1+\gamma^5)\nu^\lambda\right) \\ +\frac{ig}{2\sqrt{2}}\frac{m_e^\lambda}{M}\left(-\phi^+(\bar{\nu}^\lambda(1-\gamma^5)e^\lambda)+\phi^-(\bar{e}^\lambda(1+\gamma^5)\nu^\lambda)\right). \end{array} \tag{1.440}$$

This part of the Lagrangian is replaced by the terms

$$\begin{array}{c} -\bar{\nu}^\lambda(\gamma\partial+m_\nu^\lambda)\nu^\lambda \\ +\frac{ig}{2\sqrt{2}} W^+_\mu \left(\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5){C^{lep}}_{\lambda\kappa}e^\kappa\right) \\ +\frac{ig}{2\sqrt{2}} W^-_\mu \left(\bar{e}^\kappa {C^{lep}}^\dagger_{\kappa\lambda}\gamma^\mu(1+\gamma^5)\nu^\lambda\right) \\ +\frac{ig}{2M\sqrt{2}}\phi^+\left(-m_e^\kappa(\bar{\nu}^\lambda {C^{lep}}_{\lambda\kappa}(1-\gamma^5)e^\kappa)+m_\nu^\lambda(\bar{\nu}^\lambda {C^{lep}}_{\lambda\kappa}(1+\gamma^5)e^\kappa)\right) \\ +\frac{ig}{2M\sqrt{2}}\phi^-\left(m_e^\lambda(\bar{e}^\lambda {C^{lep}}^\dagger_{\lambda\kappa}(1+\gamma^5)\nu^\kappa)-m_\nu^\kappa(\bar{e}^\lambda {C^{lep}}^\dagger_{\lambda\kappa}(1-\gamma^5)\nu^\kappa)\right) \\ -\frac{g}{2}\frac{m_\nu^\lambda}{M}H(\bar{\nu}^\lambda\nu^\lambda)+\frac{ig}{2}\frac{m_\nu^\lambda}{M}\phi^0(\bar{\nu}^\lambda\gamma^5\nu^\lambda) \\ -\frac{1}{4}\,\bar{\nu}_\lambda\, M^R_{\lambda\kappa}\,(1-\gamma^5)\hat{\nu}_\kappa-\frac{1}{4}\,\overline{\bar{\nu}_\lambda\, M^R_{\lambda\kappa}\,(1-\gamma^5)\hat{\nu}_\kappa} \end{array} \tag{1.441}$$

while all the remaining terms of the Lagrangian of §9.4 are unchanged.
In (1.441) the matrix $C^{lep}_{\lambda\kappa}$ is the analog for leptons of the CKM mixing matrix for quarks, namely the PMNS matrix. The last term is the Majorana mass term with matrix $M^R_{\lambda\kappa}$.
Notice that, for simplicity, we used in (1.441) the same convention for leptons as the one usually adopted for quarks. Namely, we wrote the "up particles" in diagonal form and the "down particles" with the mixing matrix. In the lepton sectors the role of the "up particles" is taken by the neutrinos and that of the "down particles" by the charged leptons. The convention we follow is different from the one usually adopted in the literature on neutrino physics, where it is more natural to have the mixing matrix C^{lep} for the neutrinos, since this plays a crucial role in the theoretical description of neutrino oscillations (cf. e.g. [**233**] §11.3). It is convenient here to adopt the convention as in (1.441) for the reason that it allows us to write the Majorana mass matrix $M^R_{\lambda\kappa}$ in a simpler form.

In the following, when we write $\mathcal{L}_{SM}$ we mean the Standard Model Lagrangian with the terms (1.440) replaced by those of (1.441).

REMARK 1.115. To avoid confusing notation, we will reserve the notation $\mathcal{L}_{SM}$ and $\mathcal{L}_{Hf}$, $\mathcal{L}_{gf}$ for the terms of the full Lagrangian that include the extra terms (1.441) for mixing and Majorana masses for neutrinos, while when we want to refer to the corresponding terms (1.423) for the minimal Standard Model Lagrangian of §9.4, with the terms (1.440) instead of (1.441), we denote them by $\mathcal{L}^{\min}_{SM}$ and $\mathcal{L}^{\min}_{Hf}$, $\mathcal{L}^{\min}_{gf}$, etc.

We write here in detail the various terms of (1.423) for the case with neutrino mixing. We return to discuss each of these terms in §17 below, where we show how to recover each part of the Lagrangian from our noncommutative geometry model.

Yukawa coupling of Higgs and fermions:

(1.442)
$$\mathcal{L}_{Hf} =$$
$$\begin{gathered}
-\bar{\nu}^\lambda m_\nu^\lambda \nu^\lambda - \bar{e}^\lambda\, m_e^\lambda\, e^\lambda - \bar{u}_j^\lambda\, m_u^\lambda\, u_j^\lambda - \bar{d}_j^\lambda\, m_d^\lambda\, d_j^\lambda + \\
\tfrac{ig}{2\sqrt{2}} \tfrac{m_e^\lambda}{M} \left(-\phi^+(\bar{\nu}^\lambda(1-\gamma^5)e^\lambda) + \phi^-(\bar{e}^\lambda(1+\gamma^5)\nu^\lambda)\right) - \tfrac{g}{2}\tfrac{m_e^\lambda}{M}\left(H(\bar{e}^\lambda e^\lambda) + i\phi^0(\bar{e}^\lambda\gamma^5 e^\lambda)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^+\left(-m_d^\kappa(\bar{u}_j^\lambda C_{\lambda\kappa}(1-\gamma^5)d_j^\kappa) + m_u^\lambda(\bar{u}_j^\lambda C_{\lambda\kappa}(1+\gamma^5)d_j^\kappa)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^-\left(m_d^\lambda(\bar{d}_j^\lambda C^\dagger_{\lambda\kappa}(1+\gamma^5)u_j^\kappa) - m_u^\kappa(\bar{d}_j^\lambda C^\dagger_{\lambda\kappa}(1-\gamma^5)u_j^\kappa)\right) - \tfrac{g}{2}\tfrac{m_u^\lambda}{M}H(\bar{u}_j^\lambda u_j^\lambda) - \\
\tfrac{g}{2}\tfrac{m_d^\lambda}{M}H(\bar{d}_j^\lambda d_j^\lambda) + \tfrac{ig}{2}\tfrac{m_u^\lambda}{M}\phi^0(\bar{u}_j^\lambda\gamma^5 u_j^\lambda) - \tfrac{ig}{2}\tfrac{m_d^\lambda}{M}\phi^0(\bar{d}_j^\lambda\gamma^5 d_j^\lambda) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^+\left(-m_e^\kappa(\bar{\nu}^\lambda C^{lep}{}_{\lambda\kappa}(1-\gamma^5)e^\kappa) + m_\nu^\lambda(\bar{\nu}^\lambda C^{lep}{}_{\lambda\kappa}(1+\gamma^5)e^\kappa)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^-\left(m_e^\lambda(\bar{e}^\lambda C^{lep\dagger}{}_{\lambda\kappa}(1+\gamma^5)\nu^\kappa) - m_\nu^\kappa(\bar{e}^\lambda C^{lep\dagger}{}_{\lambda\kappa}(1-\gamma^5)\nu^\kappa)\right) - \tfrac{g}{2}\tfrac{m_\nu^\lambda}{M}H(\bar{\nu}^\lambda\nu^\lambda) + \\
\tfrac{ig}{2}\tfrac{m_\nu^\lambda}{M}\phi^0(\bar{\nu}^\lambda\gamma^5\nu^\lambda) - \tfrac{1}{4}\,\bar{\nu}_\lambda\, M^R_{\lambda\kappa}\,(1-\gamma^5)\hat{\nu}_\kappa - \tfrac{1}{4}\,\overline{\bar{\nu}_\lambda\, M^R_{\lambda\kappa}\,(1-\gamma^5)\hat{\nu}_\kappa}.
\end{gathered}$$

Notice that we have grouped the Majorana mass terms inside the $\mathcal{L}_{Hf}$ part of the Lagrangian, as well as the Dirac mass terms $-\bar{x}^\lambda m_x^\lambda x^\lambda$, for $x \in \{\nu, e, u, d\}$. The reason will be clear in §17 below where we derive the various terms of the Lagrangian.

Gauge-fermion couplings:

(1.443)
$$\mathcal{L}_{gf} =$$
$$\begin{gathered}
\tfrac{1}{2}ig_s\,\lambda_a^{ij}(\bar{q}_i^\sigma\gamma^\mu q_j^\sigma)g_\mu^a - \bar{e}^\lambda(\gamma\partial)e^\lambda - \bar{\nu}^\lambda\gamma\partial\nu^\lambda - \bar{u}_j^\lambda(\gamma\partial)u_j^\lambda - \bar{d}_j^\lambda(\gamma\partial)d_j^\lambda + \\
igs_wA_\mu\left(-(\bar{e}^\lambda\gamma^\mu e^\lambda) + \tfrac{2}{3}(\bar{u}_j^\lambda\gamma^\mu u_j^\lambda) - \tfrac{1}{3}(\bar{d}_j^\lambda\gamma^\mu d_j^\lambda)\right) + \tfrac{ig}{4c_w}Z_\mu^0\{(\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5)\nu^\lambda) + \\
(\bar{e}^\lambda\gamma^\mu(4s_w^2 - 1 - \gamma^5)e^\lambda) + (\bar{d}_j^\lambda\gamma^\mu(\tfrac{4}{3}s_w^2 - 1 - \gamma^5)d_j^\lambda) + (\bar{u}_j^\lambda\gamma^\mu(1 - \tfrac{8}{3}s_w^2 + \gamma^5)u_j^\lambda)\} + \\
\tfrac{ig}{2\sqrt{2}}W_\mu^+\left(\bar{u}_j^\lambda\gamma^\mu(1+\gamma^5)C_{\lambda\kappa}d_j^\kappa\right) + \tfrac{ig}{2\sqrt{2}}W_\mu^-\left(\bar{d}_j^\kappa C^\dagger_{\kappa\lambda}\gamma^\mu(1+\gamma^5)u_j^\lambda\right) + \\
\tfrac{ig}{2\sqrt{2}}W_\mu^+\left((\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5)C^{lep}{}_{\lambda\kappa}e^\kappa)\right) + \tfrac{ig}{2\sqrt{2}}W_\mu^-\left((\bar{e}^\kappa C^{lep\dagger}{}_{\kappa\lambda}\gamma^\mu(1+\gamma^5)\nu^\lambda)\right).
\end{gathered}$$

Notice that we have split each term of the form $-\bar{x}^\lambda(\gamma\partial + m_x^\lambda)x^\lambda$ in $\mathcal{L}_{SM}$, for $x \in \{\nu, e, u, d\}$, into a term $-\bar{x}^\lambda m_x^\lambda x^\lambda$ in $\mathcal{L}_{Hf}$ and a term $-\bar{x}^\lambda\gamma\partial x^\lambda$ in $\mathcal{L}_{gf}$.

Higgs self-coupling:

(1.444)
$$\mathcal{L}_H =$$
$$\begin{gathered}
-\tfrac{1}{2}m_h^2H^2 - \beta_h\left(\tfrac{2M^2}{g^2} + \tfrac{2M}{g}H + \tfrac{1}{2}(H^2 + \phi^0\phi^0 + 2\phi^+\phi^-)\right) + \tfrac{2M^4}{g^2}\alpha_h - \\
g\alpha_hM\left(H^3 + H\phi^0\phi^0 + 2H\phi^+\phi^-\right) - \\
\tfrac{1}{8}g^2\alpha_h\left(H^4 + (\phi^0)^4 + 4(\phi^+\phi^-)^2 + 4(\phi^0)^2\phi^+\phi^- + 4H^2\phi^+\phi^- + 2(\phi^0)^2H^2\right).
\end{gathered}$$

Self-coupling of the gauge fields:

(1.445)
$$\mathcal{L}_g =$$

$$-\tfrac{1}{2}\partial_\nu g^a_\mu \partial_\nu g^a_\mu - g_s f^{abc}\partial_\mu g^a_\nu g^b_\mu g^c_\nu - \tfrac{1}{4}g_s^2 f^{abc}f^{ade}g^b_\mu g^c_\nu g^d_\mu g^e_\nu - \partial_\nu W^+_\mu \partial_\nu W^-_\mu - M^2 W^+_\mu W^-_\mu - \tfrac{1}{2}\partial_\nu Z^0_\mu \partial_\nu Z^0_\mu - \tfrac{1}{2c_w^2}M^2 Z^0_\mu Z^0_\mu - \tfrac{1}{2}\partial_\mu A_\nu \partial_\mu A_\nu - igc_w(\partial_\nu Z^0_\mu(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - Z^0_\nu(W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + Z^0_\mu(W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) - igs_w(\partial_\nu A_\mu(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - A_\nu(W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + A_\mu(W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) - \tfrac{1}{2}g^2 W^+_\mu W^-_\mu W^+_\nu W^-_\nu + \tfrac{1}{2}g^2 W^+_\mu W^-_\nu W^+_\mu W^-_\nu + g^2 c_w^2(Z^0_\mu W^+_\mu Z^0_\nu W^-_\nu - Z^0_\mu Z^0_\mu W^+_\nu W^-_\nu) + g^2 s_w^2(A_\mu W^+_\mu A_\nu W^-_\nu - A_\mu A_\mu W^+_\nu W^-_\nu) + g^2 s_w c_w(A_\mu Z^0_\nu(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - 2A_\mu Z^0_\mu W^+_\nu W^-_\nu).$$

Notice that here we are using the Feynman gauge fixing for all the gauge fields. For the transition between the Feynman–'t Hooft gauge of the Lagrangian written as in §9.4 (as in [**285**]) and the Feynman gauge see the discussion in §9.5 above.

Minimal coupling of the Higgs fields:

$$\mathcal{L}_{Hg} = \tag{1.446}$$

$$-\tfrac{1}{2}\partial_\mu H \partial_\mu H - \partial_\mu \phi^+ \partial_\mu \phi^- - \tfrac{1}{2}\partial_\mu \phi^0 \partial_\mu \phi^0 - gMW^+_\mu W^-_\mu H - \tfrac{1}{2}g\tfrac{M}{c_w^2}Z^0_\mu Z^0_\mu H - \tfrac{1}{2}ig\left(W^+_\mu(\phi^0\partial_\mu\phi^- - \phi^-\partial_\mu\phi^0) - W^-_\mu(\phi^0\partial_\mu\phi^+ - \phi^+\partial_\mu\phi^0)\right) + \tfrac{1}{2}g\left(W^+_\mu(H\partial_\mu\phi^- - \phi^-\partial_\mu H) + W^-_\mu(H\partial_\mu\phi^+ - \phi^+\partial_\mu H)\right) + \tfrac{1}{2}g\tfrac{1}{c_w}Z^0_\mu(H\partial_\mu\phi^0 - \phi^0\partial_\mu H) - ig\tfrac{s_w^2}{c_w}MZ^0_\mu(W^+_\mu\phi^- - W^-_\mu\phi^+) + igs_wMA_\mu(W^+_\mu\phi^- - W^-_\mu\phi^+) - ig\tfrac{1-2c_w^2}{2c_w}Z^0_\mu(\phi^+\partial_\mu\phi^- - \phi^-\partial_\mu\phi^+) + igs_wA_\mu(\phi^+\partial_\mu\phi^- - \phi^-\partial_\mu\phi^+) - \tfrac{1}{4}g^2W^+_\mu W^-_\mu\left(H^2 + (\phi^0)^2 + 2\phi^+\phi^-\right) - \tfrac{1}{8}g^2\tfrac{1}{c_w^2}Z^0_\mu Z^0_\mu\left(H^2 + (\phi^0)^2 + 2(2s_w^2 - 1)^2\phi^+\phi^-\right) - \tfrac{1}{2}g^2\tfrac{s_w^2}{c_w}Z^0_\mu\phi^0(W^+_\mu\phi^- + W^-_\mu\phi^+) - \tfrac{1}{2}ig^2\tfrac{s_w^2}{c_w}Z^0_\mu H(W^+_\mu\phi^- - W^-_\mu\phi^+) + \tfrac{1}{2}g^2s_wA_\mu\phi^0(W^+_\mu\phi^- + W^-_\mu\phi^+) + \tfrac{1}{2}ig^2s_wA_\mu H(W^+_\mu\phi^- - W^-_\mu\phi^+) - g^2\tfrac{s_w}{c_w}(2c_w^2 - 1)Z^0_\mu A_\mu\phi^+\phi^- - g^2s_w^2A_\mu A_\mu\phi^+\phi^- + M\left(\tfrac{1}{c_w}Z^0_\mu\partial_\mu\phi^0 + W^+_\mu\partial_\mu\phi^- + W^-_\mu\partial_\mu\phi^+\right).$$

9.7. The Standard Model minimally coupled to gravity.

The minimal Standard Model, as well as the extension by the terms (1.441) with neutrino masses, do not include gravity. The model we are going to discuss is a further generalization, which is usually referred to as the *Standard Model minimally coupled to gravity*.

Let us first recall a few things about gravity. At the classical level, gravity is described by the theory of General Relativity. The dynamics is described by the Einstein field equation (cf. Weinberg [**298**] 7.1.13)

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda_c g_{\mu\nu} = -8\pi G\, T_{\mu\nu}, \tag{1.447}$$

where we set the speed of light to $c = 1$, $g_{\mu\nu}$ is the metric tensor with Lorentzian signature $(-,+,+,+)$, $T_{\mu\nu}$ is the energy-momentum tensor and G is Newton's gravitational constant.

The Riemann curvature tensor is given by

$$R^\lambda_{\mu\nu\kappa} = \partial_\kappa\Gamma^\lambda_{\mu\nu} - \partial_\nu\Gamma^\lambda_{\mu\kappa} + \Gamma^\rho_{\mu\nu}\Gamma^\lambda_{\kappa\rho} - \Gamma^\rho_{\mu\kappa}\Gamma^\lambda_{\nu\rho}, \tag{1.448}$$

with the Christoffel symbols

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2}g^{\lambda\kappa}\left(\partial_\mu g_{\kappa\nu} + \partial_\nu g_{\kappa\mu} - \partial_\kappa g_{\mu\nu}\right). \tag{1.449}$$

The tensors $R_{\mu\nu}$ and R are given by

$$R_{\mu\nu} = R^{\lambda}_{\mu\lambda\kappa}, \tag{1.450}$$

and

$$R = R_{\mu\nu} g^{\mu\nu}. \tag{1.451}$$

Notice that the above sign convention for the curvature tensor (1.448) is such that R is equal to $-2K$ for the 2-sphere with Gaussian curvature K. We follow [**203**] and use the term Ricci and scalar curvature for the tensors $-R_{\mu\nu}$ and $-R$.
Notice that Wald ([**289**] 4.3.21) has the opposite sign convention for the Riemann tensor with respect to the one we are following here, which is the one of [**298**].
The left-hand side of the Einstein equation (1.447) can be explained by considering the most general type of tensor

$$S_{\mu\nu} = a_1 g_{\mu\nu} + a_2 R_{\mu\nu} + a_3 R g_{\mu\nu},$$

with $a_i \in \mathbb{R}$, and imposing that it satisfies the same condition $\partial^\mu S_{\mu\nu} = 0$ as the energy-momentum tensor, for which one knows that $\partial^\mu T_{\mu\nu} = 0$.
The left-hand side of Einstein field equations (1.447) can be obtained from the Einstein–Hilbert action

$$S_{EH}(g) = \frac{1}{16\pi G} \left(\int_M R\, dv + 2\Lambda_c \int_M dv \right), \tag{1.452}$$

with $dv = \sqrt{g}\, d^4x$. This has Lagrangian density

$$\mathcal{L}_{EH} = \frac{1}{16\pi G} R + \frac{2\Lambda_c}{16\pi G}. \tag{1.453}$$

The term $\frac{2\Lambda_c}{16\pi G} \int_M dv$ is the cosmological term in the Einstein–Hilbert action.
It is important to keep track of all the sign conventions. Notice in particular that the sign in front of Λ_c is the same as that of the $R g_{\mu\nu}$ term.

Definition 1.116. *The Lagrangian density for the Standard Model minimally coupled to gravity is then simply of the form*

$$\mathcal{L} = \mathcal{L}_{EH} + \mathcal{L}_{SM}. \tag{1.454}$$

To understand why this is a good choice of Lagrangian for the minimal coupling of matter and gravity, we review briefly the argument given in §12 of [**298**].
In §9.4 and 9.6.2 we have considered the Standard Model Lagrangian (or its modification) in flat Minkowski space. One can, in fact, write the corresponding Lagrangian in any assigned Lorentzian metric $g_{\mu\nu}$. The expression becomes correspondingly more complicated, as the explicit expression of the metric tensor appears in all the inner products and pairings involved (physically, in all the raising and lowering of indices). In (1.454) one should think of the Standard Model Lagrangian as written with respect to a Lorentzian metric $g_{\mu\nu}$, for which the Einstein–Hilbert Lagrangian density is also computed.
Let S_{SM} be the action corresponding to the Lagrangian density $\mathcal{L}_{SM}$. Then the corresponding energy-momentum tensor is obtained ([**298**] §12.2) by considering the "functional derivative", i.e. the first variation δS_{SM} of the functional S_{SM} with

respect to $g_{\mu\nu}$ and writing it as a linear functional of the infinitesimal $\delta g_{\mu\nu}$ in the form

$$\delta S_{SM} = \frac{1}{2}\int T^{\mu\nu}\, \delta g_{\mu\nu}\, \sqrt{g}\, d^4x, \tag{1.455}$$

The expression (1.455) defines the energy-momentum tensor $T^{\mu\nu}$. The vanishing of δS_{SM} is a generally covariant condition, since S_{SM} is a scalar. Thus, one has the condition $\delta S_{SM} = 0$, which gives the energy-momentum conservation.
Thus, this shows that the Lagrangian density (1.454) is the correct choice that accounts for the effects of gravitation on matter and of the mass contribution of matter to the gravitational field. In fact (cf. [**298**] §12.4), the variation δS for the total action with density (1.454) takes the form

$$\delta S = \frac{1}{16\pi G}\int (R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R - \Lambda_c g^{\mu\nu} + 8\pi G T^{\mu\nu})\delta g_{\mu\nu}\, \sqrt{g}\, d^4x + \delta S_{SM},$$

so that the stationarity condition corresponds to the Einstein field equations (1.447) with the energy-momentum tensor.

It is convenient, in order to use the formalism of spectral triples in noncommutative geometry (which we introduce in §10.1 below), to work with Euclidean instead of Lorentzian signature. Here one has to be careful in spelling out precisely the conventions about signs. Passing to Euclidean signature amounts to performing a Wick rotation to imaginary time, i.e. replacing time by the change of coordinates $t \mapsto it$. What really matters as far as signs are concerned for the Euclidean action functional for gravity is that the kinetic terms have the correct positive sign to ensure that the functional is bounded below. It is well known that such positivity is spoiled by the scalar Weyl mode ([**165**]). Thus, one has to ensure that all other terms get a positive sign. To get some familiarity with this action functional we perform the simple computation of the kinetic terms explicitly (cf. [**131**] equation 4.1.4).

Let $M = \mathbb{R}^4 = V$, with the metric $g_{\mu\nu} = \delta_{\mu\nu} + \epsilon_{\mu\nu}$, where $\epsilon \in C^\infty(M, S^2V)$ is a smooth map to symmetric real 4-matrices. Let $dv = \sqrt{g}d^4x$, with $g = \det g_{\mu\nu}$.

PROPOSITION 1.117. *The second variation around flat space of the functional*

$$S(g) = \frac{1}{16\pi G}\int_M R\, dv\,, \tag{1.456}$$

is given by the quadratic form associated with the translation invariant second order operator with symbol given, up to the factor $16\pi G$*, by*

$$\langle \sigma(\xi)(\epsilon), \epsilon\rangle = \frac{1}{4}(\mathrm{Tr}(\epsilon^2) - (\mathrm{Tr}\,\epsilon)^2)\,\xi^2 + \frac{1}{2}(\mathrm{Tr}(\epsilon)\langle \epsilon\xi, \xi\rangle - \langle \epsilon\xi, \epsilon\xi\rangle) \tag{1.457}$$

The operator $\sigma(\xi)$ *(*$\xi \neq 0$*) has a four-dimensional kernel, one negative eigenvalue and its restriction to the subspace* $\{\epsilon \,|\, \mathrm{Tr}(\epsilon) = \langle \epsilon\xi, \xi\rangle\}$ *is positive.*

PROOF. For the fully covariant components of the Riemann tensor (cf. [**298**] Equation 6.6.2), one has the expression

$$\begin{aligned} R_{\lambda\mu\nu\kappa} &= \tfrac{1}{2}(g_{\lambda\nu,\kappa\mu} + g_{\kappa\mu,\lambda\nu} - g_{\mu\nu,\kappa\lambda} - g_{\kappa\lambda,\mu\nu}) \\ &+ g_{\eta\sigma}\,(\Gamma^{\eta}_{\ \nu\lambda}\Gamma^{\sigma}_{\ \mu\kappa} - \Gamma^{\eta}_{\ \kappa\lambda}\Gamma^{\sigma}_{\ \mu\nu}), \end{aligned} \tag{1.458}$$

where we use the standard abbreviation $g_{\lambda\nu,\kappa\mu} = \partial_\kappa\partial_\mu\, g_{\lambda\nu}$.

As above let $\epsilon \in C^\infty(M, S^2V)$ be a smooth map to symmetric real 4-matrices, for $M = \mathbb{R}^4 = V$. We just need to compute the terms of order two in ϵ in the expression

$$R\sqrt{g} = g^{\lambda\nu}\, g^{\mu\kappa}\, R_{\lambda\mu\nu\kappa}\, (\det g_{\mu\nu})^{1/2}$$

The term $g_{\eta\sigma}\,(\Gamma^\eta_{\nu\lambda}\Gamma^\sigma_{\mu\kappa} - \Gamma^\eta_{\kappa\lambda}\Gamma^\sigma_{\mu\nu})$ is of order two in ϵ, hence it just contributes by

$$\delta^{\lambda\nu}\,\delta^{\mu\kappa}\delta_{\eta\sigma}(\Gamma^\eta_{\nu\lambda}\Gamma^\sigma_{\mu\kappa} - \Gamma^\eta_{\kappa\lambda}\Gamma^\sigma_{\mu\nu}) = \Gamma^\eta_{\nu\nu}\Gamma^\eta_{\mu\mu} - \Gamma^\eta_{\nu\mu}\Gamma^\eta_{\nu\mu}.$$

This gives

$$\begin{aligned}&\tfrac{1}{4}\,((-\epsilon_{\nu\nu,\eta} + 2\epsilon_{\eta\nu,\nu})(-\epsilon_{\mu\mu,\eta} + 2\epsilon_{\eta\mu,\mu})\\ &-(-\epsilon_{\nu\mu,\eta} + \epsilon_{\eta\nu,\mu} + \epsilon_{\eta\mu,\nu})(-\epsilon_{\nu\mu,\eta} + \epsilon_{\eta\nu,\mu} + \epsilon_{\eta\mu,\nu})).\end{aligned}$$

To get the contribution of these terms to $\sigma(\xi)$ one just replaces each $\epsilon_{\nu\mu,\eta}$ by $\epsilon_{\nu\mu}\,\xi_\eta$. This gives

$$-\frac{3}{4}\,\mathrm{Tr}(\epsilon^2)\,\xi^2 + \frac{1}{4}\,(\mathrm{Tr}\,\epsilon)^2\,\xi^2 - \mathrm{Tr}(\epsilon)\langle\epsilon\xi,\xi\rangle + \frac{3}{2}\,\langle\epsilon\xi,\epsilon\xi\rangle. \tag{1.459}$$

The term $\frac{1}{2}\,(g_{\lambda\nu,\kappa\mu} + g_{\kappa\mu,\lambda\nu} - g_{\mu\nu,\kappa\lambda} - g_{\kappa\lambda,\mu\nu})$ is of order one in ϵ and it combines with the terms of order one in ϵ in the product $g^{\lambda\nu}\,g^{\mu\kappa}\,(\det g_{\mu\nu})^{1/2}$. These give $-\delta_{\lambda\nu}\,\epsilon_{\mu\kappa} - \epsilon_{\lambda\nu}\,\delta_{\mu\kappa} + \frac{1}{2}\,\mathrm{Tr}(\epsilon)\,\delta_{\lambda\nu}\,\delta_{\mu\kappa}$. The contributions of these terms to the symbol give

$$\mathrm{Tr}(\epsilon^2)\,\xi^2 - \frac{1}{2}\,(\mathrm{Tr}\,\epsilon)^2\,\xi^2 + \frac{3}{2}\,\mathrm{Tr}(\epsilon)\langle\epsilon\xi,\xi\rangle - 2\,\langle\epsilon\xi,\epsilon\xi\rangle. \tag{1.460}$$

Adding the contributions (1.459) and (1.460) one gets (1.457).

The kernel of $\sigma(\xi)$ is the four-dimensional space of symmetric matrices of the form

$$\epsilon = |\eta\rangle\langle\xi| + |\xi\rangle\langle\eta|,$$

with $\eta \in V$. $\square$

This shows that, after adding the gauge fixing term (cf. [**131**] equation 3.7.9)

$$\frac{1}{2}\|(\epsilon - \frac{1}{2}\mathrm{Tr}\,\epsilon)\xi\|^2, \tag{1.461}$$

which corresponds to the harmonic gauge $\Delta\, x^\mu = 0$, the restriction of the kinetic terms to the subspace of "unimodular gravity" given by the condition $\mathrm{Tr}(\epsilon) = 0$ is positive definite. Indeed, the gauge fixing term restricted to the unimodular subspace is simply $\frac{1}{2}\langle\epsilon\xi,\epsilon\xi\rangle$. It adds to the symbol $\sigma(\xi)$ to give $\frac{1}{4}\mathrm{Tr}(\epsilon^2)$ on that subspace.

Notice that positivity fails unless one adds the gauge fixing term. This suggests that one should perform the Euclidean functional integral by integrating first over the space of metrics $g_{\mu\nu}$ with a fixed volume element $\omega = \sqrt{g}\,d^4x$. The next step is then to perform the functional integration not only over all volume forms on a fixed background manifold M, but also over all manifolds. As we see below in §18, the two pieces of data given by the manifold M and the volume form ω merge together in a single datum in the framework of noncommutative geometry, where a refined incarnation of the volume form as a Hochschild cycle also contains the required information for the reconstruction of the manifold.

REMARK 1.118. Under a rescaling of the metric by a positive scalar function Ω, i.e. for the new metric $G_{\mu\nu} = \Omega^2 \, g_{\mu\nu}$, the scalar curvature transforms as

$$R(G) = \Omega^{-2} \, R(g) + 6 \, \epsilon \, \Omega^{-3} \Delta\Omega \,, \tag{1.462}$$

where the sign $\epsilon = \pm 1$ depends on the convention in the definition of the scalar curvature (one has $\epsilon = 1$ with the above convention). To obtain the physical sign of the Euclidean action of gravity ([**165**] §15.4) what matters is that the kinetic terms for Ω appear with a negative sign in the action, i.e. that one gets the term $\Omega\Delta\Omega$ (coming from $(\Omega^{-3}\Delta\Omega)\sqrt{G}$) with a *positive* sign. One then takes for g the flat metric and $\Omega = (1 + \frac{K}{4}\rho^2)^{-1}$ which (for $K > 0$) gives the round sphere. At the origin $\rho = 0$ one gets $\Delta\Omega = -2K$ and thus the physical sign of the Euclidean action gives a *negative* contribution for the sphere. This agrees with Proposition 1.117 since with our convention R is negative for the sphere.

9.8. Higher derivative terms in gravity.

It is well known that gravity as a field theory is a nonrenormalizable theory. However, at sufficiently low energies (which include the unification energy at which our model is considered) it can be treated using an "effective field theory" approach that separates out the low energy quantum corrections from the troublesome high energy terms. In the effective field theory approach (see e.g. [**122**]), one considers a Lagrangian that, in addition to the usual Einstein–Hilbert part (1.453), also involves higher derivative terms. These higher derivative terms have no effect at low energies, but become dominant at higher energies and in fact are sufficient to correct for the non-renormalizability problem, as shown in [**274**].

The first type of Euclidean higher derivative terms that are usually considered are those that come from invariant expresssions that are quadratic in the curvature, cf. [**274**]. The general form of these terms is ([**122**], [**58**])

$$\int_M \left(\frac{1}{2\eta} \, C_{\mu\nu\rho\sigma} \, C^{\mu\nu\rho\sigma} - \frac{\omega}{3\eta} \, R^2 + \frac{\theta}{\eta} \, E \right) \sqrt{g} \, d^4x. \tag{1.463}$$

Here $C_{\mu\nu\rho\sigma} \, C^{\mu\nu\rho\sigma}$ is the square of the Weyl curvature tensor and E is a topological term (cf. (1.599) (1.600) below) that integrates to the Euler characteristic

$$\chi(M) = \frac{1}{32\pi^2} \int_M E \, \sqrt{g} \, d^4x. \tag{1.464}$$

In §11.4 and §16.1 below we show that higher derivative terms appear naturally in the Lagrangian obtained from the expansion of the spectral action functional described in §11 below. We return in §17.11 below to discuss the running of the parameters η, ω and θ in (1.463) under the renormalization group equation and their role in our model of matter coupled to gravity. The theory is only an effective theory whose validity, in particular for the gravitational sector, is limited to energies up to the unification scale. Its breakdown as a fundamental theory is witnessed by the various poles of the running couplings when extended beyond the Planck mass as well as the presence of a tachyon pole in the propagator for gravitons.

9.9. Symmetries as diffeomorphisms.

We now describe a very basic reason that justifies the transition from commutative to noncommutative geometry.

An aspect of gravity that one would like to carry over to the theory of elementary particles is its *geometrization*, achieved by General Relativity. This manifests itself through the fact that the group of symmetries of gravity is given by diffeomorphisms of the underlying differentiable manifold structure of spacetime.

The scenario one can envision is a complete geometrization of the Standard Model coupled to gravity. This means turning the whole coupled theory into pure gravity on a suitable space. Now, this does not seem possible at first. The group of gauge invariance for the Einstein action $S_{EH} = \int \mathcal{L}_{EH}\sqrt{g}d^4x$ is the group $\mathrm{Diff}(M)$ of diffeomorphisms of the manifold M and the gauge invariance of the action is simply the manifestation of its geometric nature. However, the full group $\mathcal{U}$ of invariance of the action given by the Lagrangian (1.454) is richer than the group $\mathrm{Diff}(M)$ of diffeomorphisms of the manifold M, since one needs to include the group $\mathcal{G}_{SM}$ of gauge transformations of the matter sector $S_{SM} = \int \mathcal{L}_{SM}\sqrt{g}d^4x$. By construction, the group $\mathcal{G}_{SM}$ is a group of maps from M to the small gauge group $G = \mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)$. The group $\mathrm{Diff}(M)$ acts on $\mathcal{G}_{SM}$ by transformations of the base. This gives the whole group of gauge symmetries $\mathcal{U}$ of (1.454) (prior to symmetry breaking) the structure of a semidirect product

$$\mathcal{U} = \mathcal{G}_{SM} \rtimes \mathrm{Diff}(M). \tag{1.465}$$

The traditional approach to account for the structure (1.465) of the gauge symmetries, is to postulate the existence of a bundle structure over space-time M, with an action of (1.465) by bundle automorphisms. This requires an a priori distinction of certain directions in the total space as the fiber direction, with the distinction between base and fiber preserved by the symmetries. This is a natural development of the Kaluza–Klein approach to electromagnetism. Rather than following along these lines, it would seem more natural to identify a space X whose group of diffeomorphisms is directly of the form (1.465).

This search is bound to fail if one expects to find the space X among ordinary manifolds. In fact, a well known mathematical result due to W. Thurston, D. Epstein and J. Mather (cf. [**223**]) shows that the connected component of the identity in $\mathrm{Diff}(X)$ is always a simple group (see [**223**] for the precise statement). This rules out the possibility of a semidirect product structure as that of (1.465).

However, noncommutative spaces of the simplest kind readily give the answer, modulo a few subtle points. To understand what happens, notice that, for ordinary manifolds, the algebraic object corresponding to a diffeomorphism is just an automorphism $\alpha \in \mathrm{Aut}(\mathcal{A})$ of the algebra of coordinates. When an algebra is noncommutative, it admits "trivial" automorphisms, called *inner* and given by the formula

$$\alpha(x) = u\,x\,u^{-1}\,, \quad \forall x \in \mathcal{A},$$

with u an invertible element of $\mathcal{A}$. When $\mathcal{A}$ is an involutive algebra the element u is taken to be unitary (i.e. satisfying $u\,u^* = u^*\,u = 1$) so that α preserves the involution. Moreover, inner automorphisms form a subgroup

$$\mathrm{Inn}(\mathcal{A}) \subset \mathrm{Aut}(\mathcal{A})$$

which is always a normal subgroup of $\mathrm{Aut}(\mathcal{A})$.

Let us consider the simplest example, where the algebra

$$\mathcal{A} = C^\infty(M, M_n(\mathbb{C})) = C^\infty(M) \otimes M_n(\mathbb{C})$$

consists of smooth maps from a manifold M to the algebra $M_n(\mathbb{C})$ of $n \times n$ matrices. One then shows that the group $\mathrm{Inn}(\mathcal{A})$ in that case is locally isomorphic to the group $\mathcal{G}$ of smooth maps from M to the small gauge group $G = PSU(n)$ (quotient of $\mathrm{SU}(n)$ by its center) and that the general exact sequence

$$1 \to \mathrm{Inn}(\mathcal{A}) \to \mathrm{Aut}(\mathcal{A}) \to \mathrm{Out}(\mathcal{A}) \to 1$$

becomes identical to the exact sequence governing the structure of the group $\mathcal{U}$, namely

$$1 \to \mathcal{G} \to \mathcal{U} \to \mathrm{Diff}(M) \to 1.$$

It is quite striking that the terminology "internal symmetries" used in physics agrees so well with the mathematical one of "inner automorphisms". In the general case, only automorphisms that are unitarily implemented in Hilbert space will be relevant but, modulo this subtlety, one can see at once from the example above the advantage of treating noncommutative spaces on the same footing as the ordinary ones.

In the specific case of the Standard Model, we prove in §13.3, §15.1 and (1) of Theorem 1.217 below that indeed the noncommutative space that we introduce as the input data of the model will yield the correct symmetry group $\mathcal{U}$ of the form (1.465).

The next step now is to give a brief introduction to the notion of metric on a noncommutative space and the so called "spectral paradigm" of noncommutative geometry.

10. The framework of (metric) noncommutative geometry

We explained in §9.9 above why it is natural to extend the framework of geometry to "spaces" whose algebra of coordinates is no longer commutative. Such spaces appeared naturally very early in the development of quantum physics, with Heisenberg's formulation of matrix mechanics.

The extension of our basic geometric notions from ordinary manifolds to the new class of noncommutative spaces is a fundamental aspect of noncommutative geometry. An essential step is to adapt the notion of metric, i.e. the Riemannian paradigm of geometry, to the noncommutative situation. From the standpoint of physics, adopting a pragmatic point of view, we are looking for a notion of metric geometry which has the following properties.

- It contains the Riemannian paradigm $(M, g_{\mu\nu})$ as a special case.
- It does not require the commutativity of coordinates.
- It contains spaces X_z of complex dimension z suitable for the DimReg procedure.
- It provides a way of expressing the full Standard Model coupled to Einstein gravity as pure gravity on a modified space-time geometry.
- It allows for quantum corrections to the geometry.

We describe below a new paradigm which fulfills all of the above requirements. (Checking that the requirement about the Standard Model is satisfied will be our main task in the rest of this chapter.)

Since this paradigm is *spectral*, it is worthwhile to mention briefly that such spectral notions already play a key role in physics, in particular thanks to the development of the laser, in the concrete measurements of distances. In fact, the actual definition of the unit of length m in the metric system is as a specific fraction $\frac{9192631770}{299792458} \sim 30.6633...$ of the wave length of the radiation coming from the transition between two hyperfine levels of the Cesium 133 atom. Indeed the speed of light is fixed once and for all at the value of

$$c = 299792458\, m/s$$

and the second s, which is the unit of time, is defined as the time taken by 9192631770 periods of the above radiation. In adopting this prescription, one is assuming implicitly that, as a consequence of relativity, c does not depend upon the frequency of the light, a property that is the object of crucial experimental probing. The choice of Cesium is of course rather arbitrary and might eventually [**28**] be replaced by Hydrogen, which is more canonical and more abundant in the universe.
It is natural to adapt the basic paradigm of geometry to the new standard of length. We explain briefly below that this is indeed achieved by noncommutative geometry, which shows moreover how geometric spaces emerge naturally from purely spectral data.

10.1. Spectral geometry.

The Riemannian paradigm is based on the Taylor expansion in local coordinates x^μ of the square of the line element, in the form

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu \tag{1.466}$$

and the measurement of the distance $d(x, y)$ between two points is given by the geodesic formula

$$d(x, y) = \inf \int_\gamma ds \tag{1.467}$$

where the infimum is taken over all paths from x to y.

In noncommutative geometry the first basic change of paradigm has to do with the classical notion of a "real variable", which one would normally describe as a real-valued function f on a set X, i.e. as a map $f : X \to \mathbb{R}$. In fact, quantum mechanics provides a very convenient substitute. It is given by a self-adjoint operator H on Hilbert space. Notice that the choice of Hilbert space $\mathcal{H}$ is irrelevant here, since all separable infinite-dimensional Hilbert spaces are isomorphic. All the usual attributes of real variables such as their range, the number of times a real number is attained as a value of the variable, etc, have a perfect analogue in the quantum mechanical setting. The range is the spectrum of the operator H, and the spectral multiplicity $n(\lambda)$ gives the number of times a real value $\lambda \in \mathbb{R}$ is achieved.
As in the classical framework, a space X is described by the corresponding algebra $\mathcal{A}$ of coordinates, which is now concretely represented as operators on a fixed Hilbert space $\mathcal{H}$. Real coordinates are represented by self-adjoint operators and it is important to understand from the start how an ordinary "space" emerges from their joint spectrum when they happen to commute. Thus, there is no need to start from a set X as all the needed information is contained in the algebra of operators. What is surprising in the new set-up is that it gives a natural home for "infinitesimals". Indeed, it is perfectly possible for an operator to be "smaller than ϵ for any

ϵ" without being zero. This happens when the norm of the restriction of the operator to subspaces of finite codimension tends to zero when these subspaces decrease (under the natural filtration by inclusion).

Space X	Algebra $\mathcal{A}$
Real variable x^μ	Self-adjoint operator H
Infinitesimal dx	Compact operator ϵ
Integral of infinitesimal	$⨍\epsilon$ = Coefficient of $\log(\Lambda)$ in $\mathrm{Tr}_\Lambda(\epsilon)$
Line element $\sqrt{g_{\mu\nu}\, dx^\mu dx^\nu}$	ds = Fermion propagator

The corresponding operators are called "compact" and they share with naive infinitesimals all the expected algebraic properties (cf. [**68**], Chapter IV). Indeed, they form a two-sided ideal $\mathcal{K}$ of the algebra of bounded operators on $\mathcal{H}$ and the only property of the naive infinitesimal calculus that needs to be dropped is commutativity.

It is important to explain what is gained by dropping such a useful rule as commutativity. We explain this point for a specific infinitesimal, namely the "line element" ds which defines the geometry through the measurement of distances. If an infinitesimal commutes with a variable with connected range it follows that the corresponding variable x takes on a specific value. In particular with x^μ the coordinates and assuming that they commute with each other and with the line element, the latter is forced to be "localized" somewhere, which is very inconvenient. When the hypothesis of commutativity is dropped it is no longer the case that the line element ds needs to be localized and in fact it is precisely the lack of commutation of ds with the coordinates that makes it possible to measure distances.
Thus, in noncommutative geometry the basic classical formula (1.467) is replaced by the following

$$d(x,y) = \sup\{|f(x) - f(y)| \;:\; f \in \mathcal{A}\,,\ \|[D,f]\| \leq 1\}, \tag{1.468}$$

where D is the inverse of the line element ds. We explain how one can replace (1.467) with (1.468) in Proposition 1.119 below.
Notice that one should not confuse the "line element" ds with the unit of length. In the classical framework, the latter allows one to give a numerical value to the

distance between nearby points in the form (1.466). Multiplying the unit of length by a scalar λ, one divides the line element ds by λ, since ds is measured by its ratio to the unit of length.

To pass from the classical Riemannian framework to the above operator theoretic one, the key ingredient is the Dirac operator (or Atiyah–Singer operator cf. [**203**]) associated to a spin structure on a compact Riemannian manifold M. One lets $\mathcal{H} = L^2(M, S)$ be the Hilbert space of square integrable sections of the spinor bundle and one represents the algebra $\mathcal{A}$ of functions on M by simple multiplication operators

$$(f\,\xi)(x) = f(x)\,\xi(x)\,, \quad \forall f \in C^\infty(M) \text{ and } \forall \xi \in L^2(M, S).$$

One then checks directly that the computation of distances using (1.468) gives the same answer as the classical formula (1.467), in the following way.

PROPOSITION 1.119. *Let $D = \partial\!\!\!/_M$ be the Dirac operator on a Riemannian spin manifold M and let $\mathcal{H} = L^2(M, S)$ be the Hilbert space of square integrable spinors, with the algebra $C^\infty(M)$ acting on $\mathcal{H}$ as multiplication operators.*

(1) *The following conditions for $f \in C^\infty(M, \mathbb{R})$ are equivalent:*
   a) $\|[\partial\!\!\!/_M, f]\| \leq 1$
   b) $\|\nabla(f)\| \leq 1$
   c) $|f(x) - f(y)| \leq d(x, y)$, *where $d(x, y)$ is the geodesic distance* (1.467).

(2) *For complex valued functions $f \in C^\infty(M)$, one has a) $\Rightarrow$ b $\Rightarrow$ c).*

(3) *The geodesic distance is given by* (1.468).

PROOF. (1) The operator $[\partial\!\!\!/_M, f]$ is Clifford multiplication by the gradient $\nabla(f)$ of f, and the equivalence of the first two conditions follows since the operator norm is the supremum of $\|\nabla(f)\|$. Integrating the Schwartz inequality $|df| \leq \|\nabla(f)\|\|dx\|$ along a geodesic gives the implication b) $\Rightarrow$ c). The implication c) $\Rightarrow$ b) follows choosing $y^\mu = x^\mu + \epsilon\partial_\mu f$.

(2) For $a_\mu \in \mathbb{C}$ the norm of the Clifford multiplication $a = \sum a_\mu\gamma^\mu$ is the square root of the norm of $a^*a = \sum \bar{a}_\mu a_\mu + b$ where $b = \sum_{\mu\neq\nu} \bar{a}_\mu a_\nu \gamma^\mu\gamma^\nu$ has vanishing trace, thus the norm of a is larger than $\sqrt{\sum a_\mu \bar{a}_\mu}$. Applying this to $a_\mu = \partial_\mu f$ gives the first implication. Integrating the Schwartz inequality $|df| \leq \|\nabla(f)\|\|dx\|$ along a geodesic gives the implication b) $\Rightarrow$ c).
(3) One uses the function $p \mapsto d(p, x)$ to show the remaining inequality not directly implied by (2). $\square$

More generally, a refined notion of geometry (suitable in particular to deal with spaces whose coordinates do not commute) is obtained, as in the case of the distance described above, by focusing not on the traditional $g_{\mu\nu}$ but on the Dirac operator D. This way of defining a geometry by specifying the Dirac operator is meaningful both in mathematical terms (where the Dirac operator specifies the fundamental class in KO-homology) and in physics terms (where, modulo a chiral gauge transformation, the Dirac operator is the inverse of the Euclidean propagator of fermions).

The traditional notions of geometry all have natural analogs in the spectral framework. Some of these analogs are summarized in the table below and we refer to [**68**] for more details.

Geodesic equation	$\frac{d\psi(t)}{dt} = i\,\lvert D\rvert\,\psi(t)$
Geodesic Flow	$e^{i\,t\,\lvert D\rvert}$
Geodesic distance	$d(x,y) = \sup_{f\in\mathcal{A}:\lVert[D,f]\rVert\leq 1} \lvert f(x)-f(y)\rvert$
Volume form	$⨍ f\,\lvert ds\rvert^m$
Einstein action	$⨍ f\,\lvert ds\rvert^{m-2}$

10.2. Spectral triples.

A (metric) noncommutative geometry is given by a *spectral triple* $(\mathcal{A}, \mathcal{H}, D)$ in the following sense.

DEFINITION 1.120. *A spectral triple $(\mathcal{A}, \mathcal{H}, D)$ is given by an involutive unital algebra $\mathcal{A}$ represented as operators on a Hilbert space $\mathcal{H}$ and a self-adjoint operator D with compact resolvent such that all commutators $[D, a]$ are bounded for $a \in \mathcal{A}$.*

In addition to the definition above, one introduces the following notions.

DEFINITION 1.121. *A spectral triple is* even *if the Hilbert space $\mathcal{H}$ is endowed with a $\mathbb{Z}/2$- grading γ which commutes with any $a \in \mathcal{A}$ and anticommutes with D.*

DEFINITION 1.122. *A spectral triple is* finitely summable *when the resolvent of D has characteristic values $\mu_n = O(n^{-\alpha})$, for some $\alpha > 0$.*

In general, the compact operators T whose characteristic values fulfill the condition $\mu_n = O(n^{-\alpha})$ play the same role as the infinitesimals of order α in the traditional calculus.

One can introduce a first notion of dimension for spectral triples, which is the *metric dimension* and is determined by the rate of growth of the eigenvalues of the Dirac operator.

DEFINITION 1.123. *A finitely summable spectral triple is of metric dimension m if the line element $ds = D^{-1}$ is of order $1/m$.*

In this sense spectral triples that are not finitely summable represent infinite-dimensional geometries. There are other, more refined notions of dimension for spectral triples.

In fact, a more refined notion of dimension for a noncommutative geometry is not given by a single number, but by a spectrum (a subset of the complex plane), the *dimension spectrum* (cf. [**92**]). This is the subset Π of the complex plane $\mathbb{C}$ where the spectral functions (the zeta functions associated to the elements of the algebra and the Dirac operator) acquire singularities. This notion is made precise in Definition 1.133 below. We are also going to make use of another notion of dimension for noncommutative geometries, which is defined modulo 8 and is related to the KO-homology class (cf. Definition 1.124 below).

Under the hypothesis that the dimension spectrum is *simple*, i.e. that the spectral functions have at most simple poles, the residue at the pole defines a far reaching extension (cf. [**92**]) of the fundamental integral in noncommutative geometry given by the Dixmier trace (cf. [**68**]). This extends the Wodzicki residue from pseudodifferential operators on a manifold to the general framework of spectral triples, and gives meaning to $⨍ T$ in that context. It is simply given by

$$⨍ T = \mathrm{Res}_{s=0}\,\mathrm{Tr}\,(T\,|D|^{-s})\,. \tag{1.469}$$

We now recall these notions in some detail.

For our purposes, we need to discuss a further refinement of the notion of spectral triple, which is given by an additional *real structure*. This will play a fundamental role in the case of the Standard Model minimally coupled to gravity. The notion of real structure (cf. [**69**]) on a spectral triple $(\mathcal{A},\mathcal{H},D)$ is intimately related to real K-homology (cf. [**8**]) and the properties of the charge conjugation operator.

Definition 1.124. *A real structure of KO-dimension $n \in \mathbb{Z}/8$ on a spectral triple $(\mathcal{A},\mathcal{H},D)$ is an antilinear isometry $J:\mathcal{H}\to\mathcal{H}$, with the property that*

$$J^2=\varepsilon, \quad JD=\varepsilon' DJ, \quad \text{and} \quad J\gamma=\varepsilon''\gamma J \text{ (even case).} \tag{1.470}$$

The numbers $\varepsilon,\varepsilon',\varepsilon''\in\{-1,1\}$ are a function of $n \mod 8$ given by

n	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
ε	1	1	−1	−1	−1	−1	1	1
ε'	1	−1	1	1	1	−1	1	1
ε''	1		−1		1		−1	

Moreover, the action of $\mathcal{A}$ satisfies the commutation rule

$$[a,b^0]=0 \quad \forall\, a,b\in\mathcal{A}, \tag{1.471}$$

where

$$b^0 = Jb^*J^{-1} \qquad \forall b\in\mathcal{A}, \tag{1.472}$$

and the operator D satisfies the order one condition

$$[[D,a],b^0]=0 \qquad \forall\, a,b\in\mathcal{A}\,. \tag{1.473}$$

A spectral triple $(\mathcal{A},\mathcal{H},D)$ endowed with a real structure J is called a real spectral triple.

10.3. The real part of a real spectral triple.

We now show that a noncommutative geometry described by a real spectral triple $(\mathcal{A},\mathcal{H},D,J)$, in the sense of Definition 1.124 above, gives rise to an associated *commutative* real geometry $(\mathcal{A}_J,\mathcal{H},D)$, which we refer to as the *real part* of the spectral triple.

PROPOSITION 1.125. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a real spectral triple in the sense of Definition 1.124. Then the following holds.*

(1) *The equality*

$$\mathcal{A}_J = \{a \in \mathcal{A} \mid aJ = Ja\} \tag{1.474}$$

defines an involutive commutative real subalgebra of the center of $\mathcal{A}$.

(2) *$(\mathcal{A}_J, \mathcal{H}, D)$ is a real spectral triple as in Definition 1.124.*

(3) *Any $a \in \mathcal{A}_J$ commutes with the algebra generated by the sums $\sum a_i[D, b_i]$ for a_i, b_i in $\mathcal{A}$.*

PROOF. (1) By construction $\mathcal{A}_J$ is a real subalgebra of $\mathcal{A}$. Since J is isometric one has $(JaJ^{-1})^* = Ja^*J^{-1}$ for all a. Thus if $a \in \mathcal{A}_J$, one has $JaJ^{-1} = a$ and $Ja^*J^{-1} = a^*$, so that $a^* \in \mathcal{A}_J$. Let us show that $\mathcal{A}_J$ is contained in the center of $\mathcal{A}$. For $a \in \mathcal{A}_J$ and $b \in \mathcal{A}$ one has $[b, a^0] = 0$ from (1.471). But $a^0 = Ja^*J^{-1} = a^*$, so that we get $[b, a^*] = 0$.

(2) The statement follows from the fact that $\mathcal{A}_J$ is a subalgebra of $\mathcal{A}$. Notice that the result continues to hold for the complex algebra $\mathcal{A}_J \otimes_{\mathbb{R}} \mathbb{C}$ generated by $\mathcal{A}_J$.

(3) The order one condition (1.473) shows that $[D, b]$ commutes with $(a^*)^0$, hence with a since $(a^*)^0 = a$ as we saw above. $\square$

Notice that, while the real part $\mathcal{A}_J$ is contained in the center $Z(\mathcal{A})$ of $\mathcal{A}$, it can in general be smaller.

10.4. Hochschild and cyclic cohomology.

Given any (in general noncommutative) associative algebra $\mathcal{A}$ over $\mathbb{C}$, its cyclic cohomology $HC^*(\mathcal{A})$ is the cohomology of the complex (C^n_λ, b), where C^n_λ is the space of $(n+1)$-linear functionals φ on $\mathcal{A}$ such that

$$\varphi(a^1, \ldots, a^n, a^0) = (-1)^n \, \varphi(a^0, \ldots, a^n) \qquad \forall a^i \in \mathcal{A} \tag{1.475}$$

and where b is the Hochschild coboundary map given by

$$\begin{aligned} (b\varphi)\,(a^0, \ldots, a^{n+1}) \;&=\; \sum_{j=0}^{n} (-1)^j \, \varphi(a^0, \ldots, a^j a^{j+1}, \ldots, a^{n+1}) \\ &+\; (-1)^{n+1} \, \varphi(a^{n+1} a^0, \ldots, a^n). \end{aligned} \tag{1.476}$$

For $\mathcal{A} = \mathbb{C}$ one gets $HC^{2n}(\mathcal{A}) = \mathbb{C}$ and $HC^{2n+1}(\mathcal{A}) = 0$.

We refer the reader to [**68**] Chapter III.1.α for the main properties of this cohomology theory for algebras. Here we just recall briefly some basic notation which will be needed later.

The Hochschild cohomology $HH^*(\mathcal{A})$ is the cohomology of the complex (C^n, b), where one lets C^n be the linear space of $(n+1)$-multilinear forms on $\mathcal{A}$. By construction any cyclic cocycle is in particular a Hochschild cocycle and one gets a forgetful map $I : HC^* \to HH^*$.

The notions recalled in the following definition are quite useful in order to clarify the geometric significance of the basic operations, such as the map B, the cup product $\varphi \# \psi$ in cyclic cohomology, and the periodicity operator S obtained from the cup product by the generator of $HC^2(\mathbb{C})$.

DEFINITION 1.126. a) *A cycle of dimension n is a triple $\left(\Omega, d, \int\right)$, where $\Omega = \oplus_{j=0}^{n}\Omega^j$ is a graded algebra over $\mathbb{C}$, d is a graded derivation of degree 1 such that $d^2 = 0$, and $\int : \Omega^n \to \mathbb{C}$ is a closed graded trace on Ω.*

b) *Let $\mathcal{A}$ be an algebra over $\mathbb{C}$. Then a cycle over $\mathcal{A}$ is given by a cycle $\left(\Omega, d, \int\right)$ and a homomorphism $\rho : \mathcal{A} \to \Omega^0$.*

A cycle of dimension n over $\mathcal{A}$ is essentially determined by its *character*, i.e. by the $(n+1)$-linear functional τ of the form

$$\tau(a^0, \ldots, a^n) = \int \rho(a^0)\ d(\rho(a^1))\ d(\rho(a^2)) \cdots d(\rho(a^n)) \qquad \forall a^j \in \mathcal{A}. \tag{1.477}$$

Functionals obtained in this way are exactly the elements of $\operatorname{Ker} b \cap C_\lambda^n$.

Given two cycles Ω and Ω' of dimension n, their sum $\Omega \oplus \Omega'$ is defined as the direct sum of the two differential graded algebras, with $\int(\omega, \omega') = \int \omega + \int \omega'$. Given cycles Ω and Ω' of dimensions n and n', their tensor product $\Omega'' = \Omega \otimes \Omega'$ is the cycle of dimension $n + n'$, which, as a differential graded algebra, is the tensor product of (Ω, d) and (Ω', d'), with

$$\int (\omega \otimes \omega') = (-1)^{nn'} \int \omega \int \omega' \qquad \forall \omega \in \Omega\ ,\ \omega' \in \Omega'. \tag{1.478}$$

One defines in the corresponding ways the notions of direct sum and tensor product of cycles over $\mathcal{A}$ and the tensor product induces a corresponding cup product $\varphi \,\#\, \psi$ on cyclic cocycles. To express cyclic cocycles as characters of cycles one needs to use the universal differential algebra $\Omega^*(\mathcal{A})$ constructed as follows.

PROPOSITION 1.127. *Let $\mathcal{A}$ be a unital algebra over $\mathbb{C}$.*

1) *Let $\Omega^1(\mathcal{A})$ be the linear space $\mathcal{A} \otimes_{\mathbb{C}} \bar{\mathcal{A}}$, where $\bar{\mathcal{A}} = \mathcal{A}/\mathbb{C}$ is the quotient of $\mathcal{A}$ by $\mathbb{C}\,1$. Then the equalities*

$$x\,(a \otimes b)y = xa \otimes by - xab \otimes y$$

$$da = 1 \otimes a \in \Omega^1(\mathcal{A}) \quad \forall a \in \mathcal{A},$$

for any $a, b, x, y \in \mathcal{A}$, define the structure of an $\mathcal{A}$-bimodule on $\Omega^1(\mathcal{A})$ and a derivation $d : \mathcal{A} \to \Omega^1(\mathcal{A})$.

2) *Let $\mathcal{E}$ be an $\mathcal{A}$-bimodule and $\delta : \mathcal{A} \to \mathcal{E}$ a derivation with $\delta(1) = 0$. Then there exists a bimodule morphism $\rho : \Omega^1(\mathcal{A}) \to \mathcal{E}$ such that $\delta = \rho \circ d$.*

We refer the reader to [**68**] Chapter III for a proof. Proposition 1.127 shows that we can think of $(\Omega^1(\mathcal{A}), d)$ as the universal derivation of $\mathcal{A}$ in an $\mathcal{A}$-bimodule. Any element of $\Omega^1(\mathcal{A})$ is a linear combination of elements $a\,db$ for a, $b \in \mathcal{A}$ and one has $d1 = 0$.

The universal differential graded algebra $\Omega^*(\mathcal{A})$ is obtained by letting

$$\Omega^n(\mathcal{A}) = \Omega^1(\mathcal{A}) \otimes_{\mathcal{A}} \Omega^1(\mathcal{A}) \otimes_{\mathcal{A}} \cdots \otimes_{\mathcal{A}} \Omega^1(\mathcal{A})$$

be the n-fold tensor product of the bimodule $\Omega^1(\mathcal{A})$, while the differential $d : \mathcal{A} \to \Omega^1(\mathcal{A})$ extends uniquely to a square-zero graded derivation of that tensor algebra.

Any element of $\Omega^n(\mathcal{A})$ is a linear combination of elements $a^0da^1\cdots da^n$ with $a^j\in\mathcal{A}$. The product in $\Omega^*(A)$ is given by

$$\begin{aligned}(1.479)\qquad & (a^0da^1\cdots da^n)(a^{n+1}da^{n+2}\cdots da^m)\\ = & \sum_{j=1}^{n}(-1)^{n-j}a^0da^1\cdots d(a^ja^{j+1})\cdots da^nda^{n+1}\cdots da^m\\ + & (-1)^n a^0a^1da^2\cdots da^m.\end{aligned}$$

This rule follows from the requirements that $\Omega^*(\mathcal{A})$ is a right $\mathcal{A}$-module and the derivation property $d(ab)=da\,b+a\,db$.

The cyclic cocycles are exactly the characters of cycles over the algebra $\mathcal{A}$. Namely, we have the following result (cf. [**68**] Chapter III).

PROPOSITION 1.128. *Let τ be an $(n+1)$-linear functional on $\mathcal{A}$. Then the following conditions are equivalent.*

1) *There is an n-dimensional cycle $(\Omega,d,\int)$ and a homomorphism $\rho:\mathcal{A}\to\Omega^0$ such that*
$$\tau(a^0,\dots,a^n)=\int\rho(a^0)\ d(\rho(a^1))\cdots d(\rho(a^n))\qquad\forall a^0,\dots,a^n\in\mathcal{A}.$$
2) *There exists a closed graded trace $\widehat{\tau}$ of dimension n on $\Omega^*(\mathcal{A}^+)$ such that*
$$\tau(a^0,\dots,a^n)=\widehat{\tau}(a^0da^1\cdots da^n)\qquad\forall a^0,\dots,a^n\in\mathcal{A}.$$
3) *One has $\tau(a^1,\dots,a^n,a^0)=(-1)^n\ \tau(a^0,\dots,a^n)$ and*
$$\sum_{i=0}^{n}(-1)^i\ \tau(a^0,\dots,a^ia^{i+1},\dots,a^{n+1})+(-1)^{n+1}\ \tau(a^{n+1}a^0,\dots,a^n)=0$$
for any $a^0,\dots,a^{n+1}\in\mathcal{A}$.

In the second statement we denote by $\mathcal{A}^+$ the algebra of pairs (a,λ) with $a\in\mathcal{A}$, $\lambda\in\mathbb{C}$ and product
$$(a,\lambda)\cdot(a',\lambda')=(aa'+\lambda a'+\lambda' a,\lambda\lambda').$$
By construction $\mathcal{A}^+$ is unital. One needs $\mathcal{A}^+$ to handle cyclic cocycles which are not normalized, in the sense of the following definition.

DEFINITION 1.129. *A Hochschild cochain φ on a unital algebra, is* normalized *if*
$$\varphi(a^0,\dots,a^n)=0\,,\quad\text{if }\exists\, j\neq 0\,,\quad a^j=1\tag{1.480}$$
It is especially useful to consider the following example.

EXAMPLE 1.130. The simplest example is the generator of $HC^2(\mathbb{C})$. The corresponding cycle (in the sense of Definition 1.126) is given by the differential algebra $\mathcal{D}$ generated by a projection e with $e^2=e$. The two-dimensional graded trace is given by
$$\int e\,de\,de=2\pi i.\tag{1.481}$$
The map ρ from $\mathbb{C}$ to $\mathcal{D}$ defining the cycle is given by $\rho(1)=e$. The cup product of cycles induces a cup product in cyclic cohomology (cf. [**68**]) and the cup product by the generator of $HC^2(\mathbb{C})$ defines the periodicity map S.

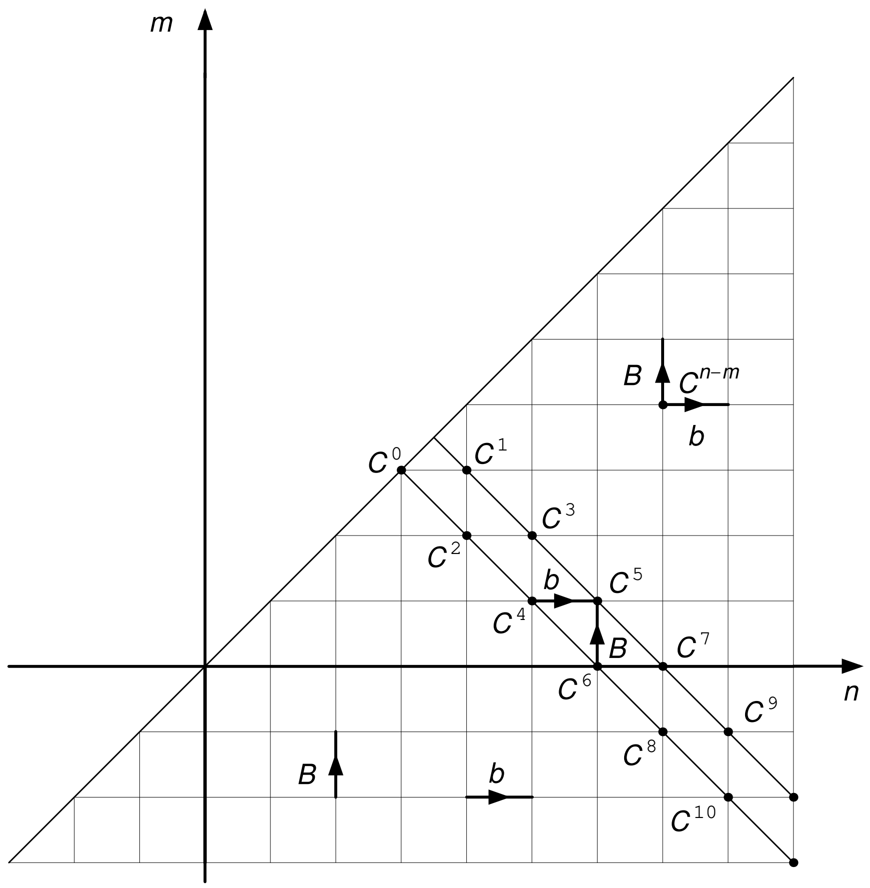


FIGURE 30. The (b, B) bicomplex

Let $\mathcal{A}$ be a unital algebra. One lets $\underline{C}^n$ be the linear space of normalized $(n+1)$-multilinear forms on $\mathcal{A}$. The Hochschild coboundary operator (1.476) preserves the normalization, namely one has

$$b : \underline{C}^n \to \underline{C}^{n+1}.$$

A normalized Hochschild cochain φ of dimension n on an algebra $\mathcal{A}$ defines (cf. [**68**]) a functional on the universal n-forms $\Omega^n(\mathcal{A})$ by the equality

$$\int_\varphi a_0\, da_1 \cdots da_n = \varphi(a_0, a_1, \ldots, a_n). \tag{1.482}$$

When φ is a Hochschild cocycle one has, using (1.479) and (1.476),

$$\int_\varphi a\, \omega = \int_\varphi \omega\, a, \quad \forall a \in \mathcal{A}. \tag{1.483}$$

The boundary operator B_0 defined by

$$B_0\, \varphi(a^0, \ldots, a^{n-1}) = \varphi(1, a^0, \ldots, a^{n-1}) - (-1)^n\, \varphi(a^0, \ldots, a^{n-1}, 1) \tag{1.484}$$

restricts on normalized cochains as

$$(B_0\varphi)(a_0,\, a_1, \cdots, a_{n-1}) = \varphi(1, a_0,\, a_1, \cdots, a_{n-1})\,. \tag{1.485}$$

It is defined in such a way that one has

$$\int_\varphi d\,\omega = \int_{B_0\varphi} \omega. \tag{1.486}$$

The operators b, B are given by (1.476) and $B = AB_0$, where

$$(A\psi)(a^0, \ldots, a^{n-1}) = \sum_0^{n-1} (-1)^{(n-1)j}\, \psi(a^j, a^{j+1}, \ldots, a^{j-1}) \tag{1.487}$$

satisfy $b^2 = B^2 = 0$ and $bB = -Bb$. Periodic cyclic cohomology, which is the inductive limit of the $HC^n(\mathcal{A})$ under the periodicity map S, admits an equivalent description as the cohomology of the (b, B) bicomplex.

We have reported here only the basic definitions and statements about cyclic cohomology without proof. We refer the reader to [**68**] and [**67**] for detailed proofs of the basic properties of cyclic cohomology, including an explanation of the meaning of the operator

$$B : HH^{n+1}(\mathcal{A}) \to HC^n(\mathcal{A})$$

in terms of the cobordism of cycles (Theorem 3.21 of [**68**]), as well as a proof of the fundamental exact sequence recalled in the following result of [**67**].

THEOREM 1.131. *The triangle*

$$\begin{tikzcd} & HH^*(\mathcal{A}) \arrow[dl, "B"'] & \\ HC^*(\mathcal{A}) \arrow[rr, "S"] & & HC^*(\mathcal{A}) \arrow[ul, "I"'] \end{tikzcd}$$

relating Hochschild and cyclic cohomologies is exact.

10.5. The local index cocycle.

We now recall briefly the local index formula in the context of noncommutative geometry [**92**]. We discuss it in the even case.

Let $(\mathcal{A}, \mathcal{H}, D)$ be a finitely summable even spectral triple. The analogue of the geodesic flow is given by the one-parameter group of automorphisms of $\mathcal{L}(\mathcal{H})$,

$$t \mapsto F_t(T) = e^{it|D|}\, T\, e^{-it|D|}\,. \tag{1.488}$$

DEFINITION 1.132. *We say that an operator T on $\mathcal{H}$ is smooth if the above map is smooth, i.e. if it belongs to $C^\infty(\mathbb{R}, \mathcal{L}(\mathcal{H}))$. We also define*

$$OP^0 := \{T \in \mathcal{L}(\mathcal{H}) \mid T \textit{ is smooth}\}\,. \tag{1.489}$$

We say that the spectral triple $(\mathcal{A}, \mathcal{H}, D)$ is regular *if it satisfies the condition*

$$a \;\; \textit{and} \;\; [D, a] \in OP^0, \quad \forall\, a \in \mathcal{A}\,. \tag{1.490}$$

Notice that the algebra $\mathcal{A}$ is not assumed to be a C^*-algebra in general i.e. it is not norm closed. At the C^*-algebra level what really matters is that the subalgebra of elements that fulfill (1.490) be norm dense.

As we already mentioned, in the context of spectral triples the usual notion of *dimension* of a space is replaced by the *dimension spectrum*. To define it one lets $\mathcal{B}$ denote the algebra generated by $\delta^k(a)$ and $\delta^k([D,a])$, for $a \in \mathcal{A}$, with

$$\delta(T) = [|D|, T] \tag{1.491}$$

the derivation that generates the geodesic flow. By construction the functions,

$$\zeta_b(z) = \mathrm{Tr}\,(b|D|^{-z}), \quad \Re(z) > m\ , \quad b \in \mathcal{B}\,, \tag{1.492}$$

are analytic. Here m denotes the crude notion of dimension given by the degree of summability of the spectral triple as in Definition 1.123.

DEFINITION 1.133. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a finitely summable regular spectral triple.*

- *The dimension spectrum is the subset Π of $\{z \in \mathbb{C}, \Re(z) \geq 0\}$ of singularities of the analytic functions $\zeta_b(z)$ for $b \in \mathcal{B}$.*
- *We say that the dimension spectrum is* simple *when the functions* (1.492) *have at most simple poles.*

In order to state the local index theorem it is important to note that one can, by a finite rank perturbation, assume the invertibility of D without changing the index pairing ([**68**] IV-1-γ). One simply replaces $(\mathcal{A}, \mathcal{H}, D)$ by $(\mathcal{A}, \tilde{\mathcal{H}}, \tilde{D})$ where

$$\tilde{\mathcal{H}} = \mathcal{H} \oplus \ker D\,, \quad \tilde{D} = D + \epsilon\, R\,, \quad \epsilon \neq 0 \tag{1.493}$$

where the action of $\mathcal{A}$ is extended to $\tilde{\mathcal{H}}$ by 0 on $\ker D$ and the finite rank operator R is 0 on the orthogonal complement of $\ker D \subset \mathcal{H}$ and is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on the remaining subspace $\ker D \oplus \ker D \subset \tilde{\mathcal{H}}$. The grading γ is extended to $\ker D$ by $-\gamma$ so that it anticommutes with $\tilde{D}$.

The local index theorem is the following result (cf. [**92**]).

THEOREM 1.134. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a regular finitely summable, even spectral triple with simple dimension spectrum. The following holds.*

- *The equality*

$$⨍ P := \mathrm{Res}_{z=0}\, \mathrm{Tr}\,(P|D|^{-z}) \tag{1.494}$$

defines a trace on the algebra generated by $\mathcal{A}$, $[D, \mathcal{A}]$ and $|D|^z$, with $z \in \mathbb{C}$.

- *Assuming $⨍ \gamma\, a = 0$ for all $a \in \mathcal{A}$, the formula*

$$\varphi_0(a) = \lim_{t \to 0} \mathrm{Tr}(\gamma\, a\, e^{-tD^2}), \quad \forall a \in \mathcal{A}, \tag{1.495}$$

defines a linear form φ_0 on $\mathcal{A}$.

- *For $n > 0$ an even integer, there is only a finite number of non-zero terms in*

$$\varphi_n(a^0, \ldots, a^n) := \sum_k c_{n,k} ⨍ \gamma\, a^0 [D, a^1]^{(k_1)} \ldots [D, a^n]^{(k_n)}\, |D|^{-n-2|k|}, \tag{1.496}$$

for all $a^j \in \mathcal{A}$. Here we are using the notation $T^{(k_i)} = \nabla^{k_i}(T)$ with $\nabla(T) = D^2T - TD^2$. The summation index k is a multi-index with $|k| = k_1 + \ldots + k_n$, and the coefficients $c_{n,k}$ are given by the formulae

$$c_{n,k} = \frac{(-1)^{|k|}}{2} \frac{\Gamma\left(|k| + n/2\right)}{k_1! \ldots k_n! \,(k_1+1)\ldots(k_1+k_2+\ldots+k_n+n)}\,. \tag{1.497}$$

- *The expression (1.496) defines the even components $(\varphi_n)_{n=0,2,\ldots}$ of a cocycle in the (b,B)-bicomplex of $\mathcal{A}$.*
- *The pairing of the cyclic cohomology class $(\varphi_n) \in HC^*(\mathcal{A})$ with $K_0(\mathcal{A})$ gives the Fredholm index of D with coefficients in $K_0(\mathcal{A})$.*

We refer the reader to [**92**] for the proof of this main theorem. In comparing the notation above and that of [**92**], notice that the trace τ_0 of Proposition II.1 of [**92**] is $\frac{1}{2}⨍$.

One of the ingredients in the proof of Theorem 1.134 will be quite useful below and we recall it here (cf. [**92**]). For any $r \in \mathbb{R}$, one lets

$$OP^r = \{T\,;\ |D|^{-r}\,T \in OP^0\}\,. \tag{1.498}$$

In general we work with operators that all make sense as acting in $\mathcal{H}^\infty = \cap\,\mathrm{Dom}\,D^N$. Also we work modulo operators of large negative order, *i.e.* $\mathrm{mod}\,OP^{-N}$ for large N. We let $\mathcal{D}(\mathcal{A})$ be the algebra generated by $\mathcal{A}$ and D considered as acting in $\mathcal{H}^\infty$. The main point is to show that one obtains an algebra by considering operators of the following form,

Definition 1.135. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a finitely summable regular spectral triple. An operator P acting in $\mathcal{H}^\infty$ is* pseudodifferential *iff for any N it can be written in the form*

$$P = T D^{-2n} \quad \mathrm{mod}\, OP^{-N}\,, \qquad T \in \mathcal{D}(\mathcal{A}) \tag{1.499}$$

The following key lemma [**92**] shows that such operators form an algebra.

Lemma 1.136. *Let $T \in OP^0$ and $n \in \mathbb{N}$.*
  (1) $\nabla^n(T) \in OP^n$
  (2) $D^{-2}\,T = \sum_0^n (-1)^k\,\nabla^k(T)\,D^{-2k-2} + R_n$, *with*

$$R_n = \,(-1)^{n+1}\,D^{-2}\,\nabla^{n+1}(T)\,D^{-2n-2}\ \in OP^{-n-3}. \tag{1.500}$$

Proof. 1) The equality

$$|D|T|D|^{-1} = T + \beta(T)\,, \quad \beta(T) = \delta(T)\,|D|^{-1}, \tag{1.501}$$

shows that for $T \in OP^0$ one has

$$D^2\,T\,D^{-2} = T + 2\,\beta\,(T) + \beta^2(T) \in OP^0. \tag{1.502}$$

Similarly one has,

$$D^{-2}\,T\,D^2 \in OP^0\,. \tag{1.503}$$

This shows that in the definition of OP^α one can put $|D|^{-\alpha}$ on either side.

To prove 1) we just need to check that $\nabla(T) \in OP^1$ and then proceed by induction. We have $\nabla(T) = D^2\,T - T\,D^2 = (D^2\,T\,D^{-2} - T)\,D^2 = (2\beta\,(T) + \beta^2(T))\,D^2 = 2\delta\,(T)\,|D| + \delta^2(T)$,

$$\nabla(T) = 2\delta\,(T)\,|D| + \delta^2(T), \tag{1.504}$$

which belongs to OP^1.
2) For $n = 0$ the statement follows from

$$D^{-2}\, T = T\, D^{-2} - D^{-2}\, \nabla(T)\, D^{-2}. \tag{1.505}$$

Next assume we have proved the result for $(n-1)$. To get it for n we must show that

$$\begin{aligned} &(-1)^n\, \nabla^n(T)\, D^{-2n-2} + (-1)^{n+1}\, D^{-2}\, \nabla^{n+1}(T)\, D^{-2n-2} \\ &= (-1)^n\, D^{-2}\, \nabla^n(T)\, D^{-2n}\,. \end{aligned} \tag{1.506}$$

Multiplying by D^{2n} on the right, with $T' = (-1)^n\, \nabla^n(T)$, we need to show that

$$T'\, D^{-2} - D^{-2}\, \nabla(T')\, D^{-2} = D^{-2}\, T',$$

which is (1.505). □

COROLLARY 1.137. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a finitely summable regular spectral triple. The pseudodifferential operators form an algebra of operators acting in $\mathcal{H}^\infty = \cap \operatorname{Dom} D^N$.*

10.6. Positivity in Hochschild cohomology and Yang-Mills action.

Let τ be a Hochschild 4-cocycle on an algebra $\mathcal{A}$. Let $\Omega^1(\mathcal{A})$ be given by Proposition 1.127. The Yang-Mills action is given by

$$YM_\tau(A) = \int_\tau (dA + A^2)^2\,, \quad \forall A \in \Omega^1(\mathcal{A}), \tag{1.507}$$

where the integration $\int_\tau$ is defined as in (1.482).
The action (1.507) is automatically gauge invariant under the gauge transformations

$$A \to \gamma_u(A) = u\, du^* + u\, A\, u^*\,, \quad \forall\, u \in \mathcal{A}\,, \quad u\, u^* = u^*\, u = 1. \tag{1.508}$$

In fact, $F(A) = dA + A^2$ transforms covariantly, that is, according to $F(\gamma_u(A)) = u\, F(A)\, u^*$.
This action functional and its precise relation with the usual Yang–Mills functional is discussed at length in [**68**] Chapter VI.
Positivity in Hochschild cohomology was defined in [**76**] for even-dimensional Hochschild cocycles $\tau \in Z^{2n}(\mathcal{A})$ as the condition

$$\int_\tau \omega\, \omega^* \geq 0\,, \quad \forall\, \omega \in \Omega^n, \tag{1.509}$$

where the adjoint ω^* is defined by

$$(a_0\, da_1 \cdots da_n)^* = (-1)^n\, da_n^* \cdots da_1^*\, a_0^*\,, \quad \forall\, a_j \in \mathcal{A}. \tag{1.510}$$

It then follows easily (cf. [**68**] Chapter VI) that the Yang-Mills action functional satisfies the following.

PROPOSITION 1.138. *Let τ be a positive Hochschild 4-cocycle on an algebra $\mathcal{A}$. Then*

$$YM_\tau(A) \geq 0\,, \quad \forall\, A \in \Omega^1\,.$$

One can find in [**68**] Chapter VI examples of positive Hochschild cocycles associated to spectral triples. We now describe a variant of these examples which plays a basic role in the computation of the variation of the spectral action under inner fluctuations as we see in §11.5 below. We let π be the representation of Ω^* given by

$$\pi(a^0 da^1 \cdots da^n) = a^0\,[D, a^1] \cdots\,[D, a^n]$$

We assume that the regular spectral triple $(\mathcal{A}, \mathcal{H}, D)$ is of dimension 4 and has simple dimension spectrum (the result extends without the last hypothesis using the Dixmier trace [**68**] instead of the residue).

LEMMA 1.139. *The equality*

$$\tau_+(a^0, a^1, a^2, a^3, a^4) = \tag{1.511}$$

$$-\,⨍\, a^0([D^2, a^1]\,[D, a^2] - [D, a^1]\,[D^2, a^2])\,([D^2, a^3]\,[D, a^4] - [D, a^3]\,[D^2, a^4])\,D^{-6}$$

defines a positive Hochschild cocycle on $\mathcal{A}$. Let F be the sign of D and let $\omega_j \in \Omega^2$. Then one has

$$\int_{\tau_+} \omega_1\,\omega_2 = ⨍\,(F\pi(\omega_1)F - \pi(\omega_1))(F\pi(\omega_2)F - \pi(\omega_2))\,D^{-4}. \tag{1.512}$$

PROOF. By construction τ_+ is a Hochschild cocycle. The right hand side of (1.512) fulfills (1.509), since for 2-forms one has $\pi(\omega^*) = (\pi(\omega))^*$. Thus, it is enough to prove (1.512). By bilinearity we can assume that $\omega_1 = a^0 da^1 da^2$ and $\omega_2 = da^3 da^4 b$. Moreover, since $[F, b]$ is of the same order as D^{-1}, we can absorb b in a^0, i.e. we assume $b = 1$. Let us show that, modulo bounded operators, we have

$$[D, \pi(\omega_1)] = a^0([D^2, a^1]\,[D, a^2] - [D, a^1]\,[D^2, a^2]).$$

Indeed, the left-hand side is

$$\begin{aligned} [D, a^0][D, a^1][D, a^2] \quad & + a^0(D[D, a^1] + [D, a^1]D)[D, a^2] \\ & - a^0[D, a^1](D[D, a^2] + [D, a^2]D) \\ = [D, a^0][D, a^1][D, a^2] \quad & + a^0([D^2, a^1]\,[D, a^2] - [D, a^1]\,[D^2, a^2]). \end{aligned}$$

Moreover, modulo bounded operators, one has

$$[D, \pi(\omega_1)] = [F, \pi(\omega_1)]\,|D| + F\,[|D|, \pi(\omega_1)] \sim [F, \pi(\omega_1)]\,|D|.$$

This gives

$$\begin{aligned} \int_{\tau_+} \omega_1\,\omega_2 = & \; -⨍[F, \pi(\omega_1)]\,|D|\,[F, \pi(\omega_2)]\,|D|\,D^{-6} \\ = & \; -⨍[F, \pi(\omega_1)]\,[F, \pi(\omega_2)]\,D^{-4}. \end{aligned}$$

Using $F^2 = 1$ one then obtains (1.512). $\square$

10.7. Cyclic cohomology and Chern-Simons action.

We now discuss briefly the invariance of the Chern-Simons action associated to a cyclic 3-cocycle on an algebra. An early instance of this action in terms of cyclic cohomology can be found in [**299**]. It is not in general invariant under gauge transformations but one has the following more subtle invariance.

PROPOSITION 1.140. *Let ψ be a cyclic 3-cocycle on $\mathcal{A}$. The functional*

$$CS_\psi(A) = \int_\psi A\, d\, A + \frac{2}{3}\, A^3\,, \quad A \in \Omega^1 \tag{1.513}$$

transforms under the gauge action $\gamma_u(A) = u\, du^* + u\, A\, u^*$ as

$$CS_\psi(\gamma_u(A)) = CS_\psi(A) + \frac{1}{3}\, \langle \psi, u \rangle, \tag{1.514}$$

where $\langle \psi, u \rangle$ is the pairing between $HC^3(\mathcal{A})$ and $K_1(\mathcal{A})$.

PROOF. Let $A' = \gamma_u(A) = u\, du^* + u\, A\, u^*$. One has

$$dA' = du\, du^* + du\, A\, u^* + u\, dA\, u^* - u\, A\, du^*,$$

$$\begin{aligned} A'\, dA' = {} & u\, du^*\, du\, du^* + u\, du^*\, du\, A\, u^* + u\, du^*\, u\, dA\, u^* - u\, du^*\, u\, A\, du^* \\ & + u\, A\, u^*\, du\, du^* + u\, A\, u^*\, du\, A\, u^* + u\, A\, dA\, u^* - u\, A^2\, du^*. \end{aligned}$$

Thus, using the graded trace property of $\int_\psi$, one gets

$$\int_\psi (\, A'\, dA' - A\, dA) =$$

$$\begin{aligned} \int_\psi \quad & (u\, du^*\, du\, du^* + du^*\, du\, A - du^*\, u\, du^*\, u\, A \\ & + u^*\, du\, du^*\, u\, A + du^*\, u\, dA + u^*\, du\, A^2 - du^*\, u\, A^2). \end{aligned}$$

Using

$$\int_\psi du^*\, u\, dA = - \int_\psi du^*\, du\, A,$$

this gives

$$\int_\psi (\, A'\, dA' - A\, dA) = \int_\psi (u\, du^*\, du\, du^* + 2\, u^*\, du\, du^*\, u\, A + 2\, u^*\, du\, A^2).$$

Next one has

$$\int_\psi (A'^3 - A^3) = \int_\psi ((u\, du^*)^3 + 3\, (u\, du^*)^2\, u\, A\, u^* + 3\, u\, du^*\, u\, A^2\, u^*).$$

Since $du^*\, u = -u^*\, du$, the terms in A^2 cancel in the variation of CS_ψ. Similarly, one has $du^*\, u\, du^*\, u = -u^*\, du\, du^*\, u$ so that the terms in A also cancel. Thus, one obtains

$$CS_\psi(\gamma_u(A)) - CS_\psi(A) = \int_\psi (u\, du^*\, du\, du^* + \frac{2}{3}(u\, du^*)^3).$$

One has $(u\, du^*)^3 = -u\, du^*\, du\, du^*$ which gives the required result. $\square$

10.8. Inner fluctuations of the metric.

Exactly as the inner automorphisms of a noncommutative space correspond to the internal symmetries of physics (see §9.9), the metric of a noncommutative space admits natural inner fluctuations. These give rise to a natural foliation of the space of metrics. In the case of the Standard Model minimally coupled to gravity, we show in §15.2 and §15.4 below that the inner fluctuations of noncommutative geometry correspond to the gauge bosons other than the graviton.

The inner fluctuations in noncommutative geometry are generated by the existence of Morita equivalences (cf. [**253**]). Given an algebra $\mathcal{A}$, a Morita equivalent algebra $\mathcal{B}$ is the algebra of endomorphisms of a finite projective (right) module $\mathcal{E}$ over $\mathcal{A}$,

$$\mathcal{B} = \mathrm{End}_{\mathcal{A}}(\mathcal{E}) \tag{1.515}$$

If $\mathcal{A}$ acts on the Hilbert space $\mathcal{H}$, then $\mathcal{B}$ acts in a natural manner on the tensor product

$$\mathcal{H}' = \mathcal{E} \otimes_{\mathcal{A}} \mathcal{H}, \tag{1.516}$$

which is a Hilbert space, provided that $\mathcal{E}$ is endowed with a Hermitian structure given by an $\mathcal{A}$-valued positive sesquilinear form on $\mathcal{E}$ satisfying the same rules as an inner product.

Defining the analogue D' of the operator D for $(\mathcal{B}, \mathcal{H}')$ requires the choice of a *Hermitian connection* ∇ on $\mathcal{E}$. In fact, one would at first think of replacing D by the operator

$$D'(\xi \otimes \eta) = \xi \otimes D\eta,$$

but this is not compatible with the tensor product over $\mathcal{A}$, since in general the operator D does not commute with elements of $\mathcal{A}$. One needs to introduce a connection to correct for this problem.

A connection is a linear map $\nabla : \mathcal{E} \to \mathcal{E} \otimes_{\mathcal{A}} \Omega^1_D$ satisfying the Leibniz rule

$$\nabla(\xi a) = (\nabla\xi)a + \xi \otimes da\,, \quad \forall\, \xi \in \mathcal{E}\ ,\ a \in \mathcal{A},$$

with $da = [D, a]$ and with

$$\Omega^1_D = \{\sum_j a_j\,[D, b_j] \,|\, a_j,\, b_j \in \mathcal{A}\}. \tag{1.517}$$

Notice that Ω^1_D is by construction a bimodule over $\mathcal{A}$. One then defines the new Dirac operator D' as

$$D'(\xi \otimes \eta) = \xi \otimes D\eta + (\nabla\xi)\,\eta. \tag{1.518}$$

This combination is well defined as an operator on $\mathcal{H}' = \mathcal{E} \otimes_{\mathcal{A}} \mathcal{H}$.

Any algebra $\mathcal{A}$ is in particular Morita equivalent to itself (with $\mathcal{E} = \mathcal{A}$). When one applies the construction of the new Dirac operator D' of (1.518) in this special case, one gets the inner deformations of the spectral geometry. These replace the operator D by

$$D \mapsto D + A \tag{1.519}$$

where $A = A^*$ is an arbitrary selfadjoint element of Ω^1_D.

This discussion on the inner fluctuations of the metric adapts to the presence of the additional structure given by the charge conjugation operator (real structure) as in Definition 1.124. A key role of the real structure J is in defining the *adjoint action* of the unitary group $\mathcal{U}$ of the algebra $\mathcal{A}$ on the Hilbert space $\mathcal{H}$. In fact, one defines a right $\mathcal{A}$-module structure on $\mathcal{H}$ by

$$\xi\, b = b^0\, \xi\,, \quad \forall\, \xi \in \mathcal{H}\,, \quad b \in \mathcal{A} \tag{1.520}$$

The unitary group of the algebra $\mathcal{A}$ then acts by the "adjoint representation" in $\mathcal{H}$ in the form

$$\xi \in \mathcal{H} \to \mathrm{Ad}(u)\,\xi = u\xi u^*\,, \quad \forall\, \xi \in \mathcal{H}\,, \quad u \in \mathcal{A}\,, \quad uu^* = u^*u = 1\,, \tag{1.521}$$

and the perturbation (1.519) gets replaced by

$$D \mapsto D + A + \varepsilon'\, JAJ^{-1}, \tag{1.522}$$

where ε' is specified by the table of Definition 1.124.
The following result describes the relation between the adjoint action of the unitary group of the algebra $\mathcal{A}$ and the gauge potentials $A \in \Omega^1_D$ with $A = A^*$.

PROPOSITION 1.141. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a real spectral triple with antilinear isometry J fulfilling* (1.472) *and* (1.473). *Then for any gauge potential $A \in \Omega^1_D$ with $A = A^*$ and any unitary $u \in \mathcal{A}$, one has*

$$\mathrm{Ad}(u)(D + A + \varepsilon'\, JAJ^{-1})\mathrm{Ad}(u^*) = D + \gamma_u(A) + \varepsilon'\, J\gamma_u(A)J^{-1},$$

where

$$\gamma_u(A) = \, u\,[D, u^*] + \, u\,A\,u^*.$$

PROOF. Let $U = \mathrm{Ad}(u)$. One has $U = u\,v = v\,u$, where $v = JuJ^{-1}$. Thus,

$$\begin{aligned} UDU^* &= \; u(vDv^*)u^* = u(D + v[D, v^*])u^* \\ &= \; uDu^* + v[D, v^*] = D + u[D, u^*] + v[D, v^*], \end{aligned}$$

where we used (1.472) and (1.473). We have

$$v[D, v^*] = JuJ^{-1}[D, Ju^*J^{-1}] = \varepsilon' J\, u[D, u^*]\, J^{-1}.$$

Similarly, we have $UAU^* = uAu^*$ and $UJ\,A\,J^{-1}U^* = J\,uAu^*\,J^{-1}$. □

The next proposition shows that inner fluctuations of inner fluctuations are inner fluctuations.

PROPOSITION 1.142. *1) Let $(\mathcal{A}, \mathcal{H}, D)$ be a spectral triple and let $D' = D + A$, for some $A \in \Omega^1_D$ with $A = A^*$. Then, for any $B \in \Omega^1_{D'}$ with $B = B^*$, one has*

$$D' + B = \, D + A'\,, \quad A' = A + B \in \Omega^1_D \tag{1.523}$$

2) Let $(\mathcal{A}, \mathcal{H}, D)$ be a real spectral triple with antilinear isometry J fulfilling (1.472) *and* (1.473). *Let $A \in \Omega^1_D$ with $A = A^*$ and let $D' = D + A + \varepsilon'\, JAJ^{-1}$. Then, for any $B \in \Omega^1_{D'}$ with $B = B^*$, one has*

$$D' + B + \varepsilon'\, JBJ^{-1} = \, D + A' + \varepsilon'\, JA'J^{-1}\,, \quad A' = A + B \in \Omega^1_D \tag{1.524}$$

PROOF. To check 1) it is enough to show that $\Omega^1_{D'} \subset \Omega^1_D$. By construction, elements of $\Omega^1_{D'}$ are linear combinations of $a[D', b]$, for $a,\, b \in \mathcal{A}$. In fact, we have $a[D', b] = a[D, b] + a[A, b] \in \Omega^1_D$, since Ω^1_D is a bimodule over $\mathcal{A}$.
Similarly, to check 2) one has to show that $\Omega^1_{D'} \subset \Omega^1_D$. For $a,\, b \in \mathcal{A}$ one has $a[D', b] = a[D, b] + a[A, b] + \varepsilon'\, a[JAJ^{-1}, b]$, but then we have $[JAJ^{-1}, b] = 0$, using (1.472) and (1.473). Thus, the inclusion $\Omega^1_{D'} \subset \Omega^1_D$ follows as above. □

It is important to consider the inner fluctuations $D \mapsto D + A + \varepsilon' JAJ^{-1}$ as being obtained in two stages

(1) $D \mapsto D + A$
(2) $D + A \mapsto D + A + \varepsilon' JAJ^{-1}$

In fact, the resulting operator $D + A + \varepsilon' JAJ^{-1}$ is the most general operator D' obtained from inner fluctuations relative to $\mathcal{A}$ and then relative to $\mathcal{A}^0 = J\mathcal{A}J^{-1}$, with the property that D' fulfills the same commutation relation

$$JD'J^{-1} = \varepsilon' D'$$

with J as the original D. In this way one can extend many results valid for inner fluctuations of the form $D \mapsto D + A$ to the more sophisticated ones of the form (1.522).

Remark 1.143. Notice that, although transitive, the relation "D' is an inner fluctuation of D" is not symmetric. In fact, one can construct simple examples where the operator 0 is obtained as inner fluctuation of a non-zero D. This relation is similar to the relation obtained from a singular foliation, such as the foliation associated to a flow with singularities given by the fixed points. We see in §15.2 below a concrete example of such a singularity, in the case of the Higgs mechanism in the Standard Model, cf. Remark 1.203.

11. The spectral action principle

We give in this section an account of the joint work of the first author with A. Chamseddine on the spectral action principle [**45**], [**46**], [**47**], [**48**]. This will be the main tool we use in the following sections to recover the Lagrangian of the Standard Model minimally coupled to gravity from a simple noncommutative geometry.

The starting point to motivate the spectral action is a discussion of observables in gravity. By the principle of gauge invariance, the only quantities which have a chance to be observables in gravity are those which are invariant under the gauge group i.e. the group of diffeomorphisms of the space-time M. Assuming first that we deal with a classical manifold, one can form a number of such invariants (under suitable convergence conditions) as integrals of the form

$$\int_M F(K)\,\sqrt{g}\,d^4x \tag{1.525}$$

where $F(K)$ is a scalar invariant function of the Riemann curvature K.

The scalar curvature is one example of such a function but there are many others. We refer to [**143**] for other more complicated examples of such invariants, where those of the form (1.525) appear as the *single integral* observables, i.e. those which add up when evaluated on the direct sum of geometric spaces.

While in principle a quantity like (1.525) is observable, it is in practice almost impossible to evaluate, since it involves the knowledge of the entire space-time and is in that way highly non-localized. On the other hand, spectral data are available in localized form anywhere and are (asymptotically) of the form (1.525) when they are of the additive form

$$\mathrm{Tr}(f(D/\Lambda)), \tag{1.526}$$

where D is the Dirac operator and f is a positive even function of the real variable while the parameter Λ fixes the mass scale. The data of spectral lines are intimately related to the Dirac Hamiltonian, hence to the geometry of "space".

The *spectral action principle* asserts that (1.526) is the fundamental action functional S that can be used both at the classical level to compare different geometric spaces and at the quantum level in the functional integral formulation, after Wick rotation to Euclidean signature.

In the following, we work under the assumption that we have an asymptotic expansion

$$\mathrm{Trace}\,(e^{-tD^2}) \sim \sum a_\alpha\, t^\alpha \qquad (t \to 0) \tag{1.527}$$

The absence of a $\log(t)$ in the expansion corresponds to the dimension spectrum being *simple*. Notice that this assumption is not always satisfied. For instance, it is known to fail in the case of manifolds with conical singularities (cf. [**206**]).

Under the above assumption, the detailed form of the even function f in (1.526) is largely irrelevant because of the known simple relation between the heat expansion and residues which we now recall ([**303**]). The ζ-function is defined by

$$\zeta_D(s) = \mathrm{Tr}\,(|D|^{-s}) = \mathrm{Tr}\,(\Delta^{-s/2})\,, \quad \Delta = D^2\,. \tag{1.528}$$

This only makes sense if D is invertible but this hypothesis is harmless since one can replace D by $D + \epsilon P$ where P is the orthogonal projection on $\ker D$ to make it invertible and then eliminate ϵ by taking the limit when $\epsilon \to 0$. Equivalently one restricts to the orthogonal complement of the finite dimensional subspace $\ker D \subset \mathcal{H}$ so that (1.528) makes sense. The relation between the asymptotic expansion (1.527) and the ζ-function is given by:

LEMMA 1.144.
- *A non-zero term a_α with $\alpha < 0$ gives a pole of ζ_D at -2α with*

$$\mathrm{Res}_{s=-2\alpha}\, \zeta_D(s) = \frac{2\, a_\alpha}{\Gamma(-\alpha)} \tag{1.529}$$

- *The absence of* $\log t$ *terms gives regularity at* 0 *for* ζ_D *with*

$$\zeta_D(0) + \dim\ker D = a_0\,. \tag{1.530}$$

PROOF. For the positive operator $\Delta = D^2$ one has

$$|D|^{-s} = \Delta^{-s/2} = \frac{1}{\Gamma\left(\frac{s}{2}\right)} \int_0^\infty e^{-t\Delta}\, t^{s/2-1}\, dt. \tag{1.531}$$

Replacing $\mathrm{Trace}\,(e^{-t\Delta})$ by $a_\alpha\, t^\alpha$ and using

$$\int_0^1 t^{\alpha+s/2-1}\, dt = (\alpha + s/2)^{-1}$$

one gets the first statement. The second follows from the equivalence

$$\frac{1}{\Gamma\left(\frac{s}{2}\right)} \sim \frac{s}{2}\,, \quad s \to 0$$

so that only the pole part at $s = 0$ of

$$\int_0^\infty \mathrm{Tr}(e^{-t\Delta})\, t^{s/2-1}\, dt$$

contributes to the value $\zeta_D(0)$. This pole part is given by

$$a_0 \int_0^1 t^{s/2-1}\, dt = a_0 \frac{2}{s}$$

so that one gets (1.530). This holds when D is invertible and in the general case the addition of ϵP amounts to the replacement

$$\zeta_D(0) \mapsto \zeta_D(0) + \dim\ker D\,.$$

$\square$

THEOREM 1.145. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a spectral triple fulfilling* (1.527). *Then the spectral action* (1.526) *can be expanded in powers of the scale Λ in the form*

$$\mathrm{Tr}(f(D/\Lambda)) \sim \sum_{\beta\in\Pi} f_\beta\, \Lambda^\beta \,⨍ |D|^{-\beta} + f(0)\,\zeta_D(0) + \cdots, \tag{1.532}$$

with the summation over the dimension spectrum Π. Here the function f only appears through the scalars

$$f_\beta = \int_0^\infty f(v)\, v^{\beta-1}\, dv. \tag{1.533}$$

The terms involving negative powers of Λ involve the full Taylor expansion of f at 0.

PROOF. Consider a test function $k(u)$ which is given as a simple superposition of exponentials as a Laplace transform

$$k(u) = \int_0^\infty e^{-su}\, h(s)\, ds. \tag{1.534}$$

We can write formally

$$k(t\Delta) = \int_0^\infty e^{-st\Delta}\, h(s)\, ds \tag{1.535}$$

and

$$\mathrm{Trace}\,(k(t\Delta)) \sim \sum a_\alpha\, t^\alpha \int_0^\infty s^\alpha\, h(s)\, ds\,. \tag{1.536}$$

For $\alpha < 0$ one has,

$$s^\alpha = \frac{1}{\Gamma(-\alpha)} \int_0^\infty e^{-sv}\, v^{-\alpha-1}\, dv$$

and

$$\int_0^\infty s^\alpha\, h(s)\, ds = \frac{1}{\Gamma(-\alpha)} \int_0^\infty k(v)\, v^{-\alpha-1}\, dv$$

so that

$$\begin{aligned} \mathrm{Trace}\,(k(t\Delta)) \;\sim\; & \sum_{\alpha<0} \frac{1}{2}\, \mathrm{Res}_{s=-2\alpha}\, \zeta_D(s) \int_0^\infty k(v)\, v^{-\alpha-1}\, dv\; t^\alpha \\ & + \; \zeta_D(0)\, k(0) + \ldots \end{aligned} \tag{1.537}$$

Thus, with f even as in (1.532), we write $f(u) = k(u^2)$ and obtain (1.532) using

$$\frac{1}{2} \int_0^\infty k(v)\, v^{\beta/2-1}\, dv = \int_0^\infty f(u)\, u^{\beta-1}\, du,$$

for $\beta = -2\alpha \in \Pi$. $\square$

The above gives a formal proof of the asymptotic expansion of the spectral action functional. For a more detailed analysis of the actual expansion, we refer the reader to the book [**132**].

The term in Λ^4 in the spectral action (1.532) is proportional to $⨍ ds^4$ which, in the 4-dimensional case, gives a cosmological term. As we see later, the natural constraint in the set-up of the functional integral will provide a homological meaning to this term cf. equation (1.874) below.

The term in Λ^2 in the spectral action (1.532) is proportional to $⨍ ds^2$ which, in the usual 4-dimensional Riemannian case (cf. [**179**]), gives the Einstein-Hilbert action functional with the physical sign for the Euclidean functional integral, provided the moment $f_2 > 0$ (which is the case if f is a positive function). This term will be analysed in the general case in §11.1 below.

The term in Λ^0 in the spectral action (1.532) yields the Yang–Mills action for the gauge fields corresponding to the internal degrees of freedom of the metric. This result will be obtained first by a direct computation in §11.4 using the Seeley–DeWitt coefficients 11.2. We then explain (§11.5) in the general framework that under a simple hypothesis of vanishing of the tadpole (which holds in the usual Riemannian case) the variation under inner fluctuations of the scale-independent part of the spectral action can be expressed as the sum of a Yang–Mills action and a Chern–Simons action.

In particular, it is convenient to introduce the following notation.

Definition 1.146. *The tadpole term of order Λ^α is the term linear in $A \in \Omega^1_D$ in the Λ^α part of the expansion* (1.532) *of the spectral action.*

The vanishing of all these tadpole terms indicates that a given geometry $(\mathcal{A}, \mathcal{H}, D)$ is a critical point for the spectral action. In general, one can impose weaker conditions of vanishing of tadpole terms, as we will see in the case of the Higgs mechanism in the Standard Model.

11.1. Terms in Λ^2 in the spectral action and scalar curvature.

Throughout this section we let $(\mathcal{A}, \mathcal{H}, D)$ be a regular spectral triple with simple dimension spectrum and with metric dimension 4.

Definition 1.147. *The* scalar curvature *of $(\mathcal{A}, \mathcal{H}, D)$ is the functional given by*

$$\mathcal{R}(a) = ⨍ a\, D^{-2}\,, \quad \forall a \in \mathcal{A}\,. \tag{1.538}$$

The terminology is motivated by the following result of [**179**].

Theorem 1.148. *In the case of four-dimensional Riemannian spin manifolds M, the functional $\mathcal{R}$ is given by*

$$\mathcal{R}(f) = -\frac{1}{24\,\pi^2} \int f(x)\, s(x)\, \sqrt{g}\, d^4x \tag{1.539}$$

where $s = -R$ is the scalar curvature.

PROOF. By definition the scalar curvature s at a point $x \in M$ is, up to sign, twice the *trace* of the Riemann curvature tensor viewed as an endomorphism of $\wedge^2 T_x(M)$ and is *positive* for the sphere. (With the sign convention for the curvature tensor used by Milnor [**229**] p. 51, one gets a plus sign.) When acting on 1-forms the formula for the curvature is $\nabla_X \nabla_Y \omega - \nabla_Y \nabla_X \omega - \nabla_{[X,Y]} \omega$, cf. Wald [**289**] p. 36-40.

Our chosen normalization for the Riemann curvature tensor (1.448) is the same as Gilkey ([**144**] §2.3) and Lawson-Michelsohn ([**203**], Proposition 4.6). It reads

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{1.540}$$

Following Lawson–Michelsohn [**203**], both the Ricci curvature and the scalar curvature have the usual sign (positive for spheres), thanks to the minus signs in equations 8.10 and 8.17 of [**203**]. We follow the same convention and we define the scalar curvature by setting

$$s = -R. \tag{1.541}$$

As already mentioned, Gilkey [**144**] uses the same convention (1.540) for R as Lawson-Michelsohn, but defines the scalar curvature as $\frac{1}{2}$ of the above s. With our convention, s is the term which appears in the Lichnerowicz formula (cf. [**203**] Theorem 8.8)

$$D^2 = \nabla^* \nabla + \frac{1}{4}\, s \tag{1.542}$$

The formula (1.542) decomposes the operator $P = D^2$ in the canonical form of [**144**] Lemma 4.8.1. The operators of [**144**] are $P_\nabla = \nabla^* \nabla$ and $E = -\frac{1}{4}\, s$, where s is as in (1.541) above.

By Theorem 4.8.16 of [**144**] the coefficient a_2 in the heat expansion of P is given by

$$a_2(x, D^2) = (4\pi)^{-2} \mathrm{Tr} \left(-\frac{1}{6} R_{ijij} + E \right), \tag{1.543}$$

where the trace is taken in the spinor space, hence it multiplies the scalar value by 4. With the notations of [**144**], one has $R_{ijij} = -s$, so that (1.543) gives in front of s the coefficient

$$(4\pi)^{-2} \times 4 \times (\frac{1}{6} - \frac{1}{4}) = -\frac{1}{48\,\pi^2}.$$

By (1.529) there is an additional factor of 2 in the relation with the residue, and we get (1.539). ☐

In particular, we see that, when written with the notation (1.451), equation (1.539) becomes of the form

$$\mathcal{R}(f) = \frac{1}{24\,\pi^2} \int f(x)\, R(x)\, \sqrt{g}\, d^4 x, \tag{1.544}$$

which has the correct sign for the Euclidean functional integral as explained in Proposition 1.117.

Notice that, if one had chosen the *scalar* Laplacian instead of the square of the Dirac operator, one would have obtained the wrong sign in front of the scalar curvature.

We now investigate, in the general case, the stability of the scalar curvature as defined by (1.538) under inner fluctuations of the metric of the form $D \mapsto D + A$

with $A = A^* \in \Omega^1_D$. The coefficient of Λ^2, given by $⨍(D+A)^{-2}$, has a priori an A dependence. If we write $(D+A)^2 = D^2 + X$, we obtain

$$⨍ (D+A)^{-2} - ⨍ D^{-2} = - ⨍ D^{-2}\, X\, D^{-2} + ⨍ D^{-2}\, X\, D^{-2}\, X\, D^{-2}, \tag{1.545}$$

where $X = DA + AD + A^2$ has order one, so that the other terms vanish. We can rewrite this as

$$- ⨍ (DA + AD)\, D^{-4} + ⨍ (DA+AD)^2\, D^{-6} - ⨍ A^2\, D^{-4}\,. \tag{1.546}$$

We now analyse these various terms.
The first term we meet in (1.546) with a coefficient -2 is

$$⨍ A\, D^{-3}\,. \tag{1.547}$$

This is the Λ^2 dependent *linear* piece in the spectral action, i.e. the tadpole term of order Λ^2 in the sense of Definition 1.146.

LEMMA 1.149. *The term* (1.547) *vanishes in the case of a Riemannian spin manifold* M.

PROOF. In terms of local coordinates x^μ the Dirac operator is of the form

$$D = \partial\!\!\!/_M = \sqrt{-1}\,\gamma^\mu\, \nabla^s_\mu$$

where ∇^s is the spin connection. The elements of Ω^1_D are the sections of the bundle $\mathrm{End}(S)$ coming from Clifford multiplication $\gamma(\omega)$ by 1-forms $\omega \in C^\infty(M, T^*M)$. Thus A is of the form $A = \gamma(\omega)$. The total symbol of the operator D^{-3} is of the form

$$\sigma(D^{-3})(x,\xi) = \sigma_{-3}(x,\xi) + \sigma_{-4}(x,\xi) + \cdots$$

where $\sigma_{-n}(x,\xi) \in \mathrm{End}(S_x)$ is an endomorphism of the fiber S_x of the spinor bundle. The total symbol of a product of pseudodifferential operators is given by (cf. [**144**] §1.2)

$$\sigma^{PQ}(x,\xi) = \sum \frac{(-i)^{|\alpha|}}{\alpha!} \partial^\alpha_\xi \sigma^P(x,\xi)\, \partial^\alpha_x \sigma^Q(x,\xi)\,. \tag{1.548}$$

This shows that the symbol of order -4 of $A\,D^{-3}$ is $\gamma(\omega)\sigma_{-4}(x,\xi)$ and the residue is of the form

$$⨍ A\, D^{-3} = \int_M \mathrm{tr}(\gamma(\omega) b(x))\, dv$$

where $b(x) \in \mathrm{End}(S_x)$ is obtained by integrating $\sigma_{-4}(x,\xi)$ over the unit sphere of the cotangent space $T^*_x(M)$ and dv is the Riemannian volume form on M. By construction one has $b \in C^\infty(M, \mathrm{End}(S))$ and to compute $b(x)$ one can use normal coordinates at x. In these coordinates the total symbol of D at x is $\sqrt{-1}\,\gamma^\mu\,\xi_\mu$ and the symbol of order -5 of D^{-4} vanishes (cf. [**179**]). Thus using (1.548) one gets $\sigma_{-4}(x,\xi) = 0$ and $b(x) = 0$. □

Since (1.547) is the Λ^2 dependent *linear* piece in the spectral action, if we expand at an extremum of the spectral action which is independent of Λ, it is natural to assume that (1.547) vanishes, namely that we have

$$⨍ AD^{-3} = 0 \qquad \forall\, A \in \Omega^1_D\,. \tag{1.549}$$

There are however very interesting situations, such as the Higgs case, where the extremum depends upon Λ.
Under the hypothesis of vanishing of the order Λ^2 tadpole term, we have the following result.

LEMMA 1.150. *If (1.549) holds, then one has*

$$\fint [D,a][D,b]\,ds^4 = -\fint a\,[D^2,b]\,ds^4 \qquad \forall\, a,b \in \mathcal{A}\,.$$

PROOF. By (1.549) and using the identity

$$D\,a\,[D,b] + a\,[D,b]\,D = [D,a][D,b] + a\,[D^2,b]$$

one obtains

$$\fint (D\,a\,[D,b] + a\,[D,b]\,D)\,D^{-4} = 0.$$

□

Let us now assume also the condition that the functional playing the role of the scalar curvature is a trace, namely

$$\mathcal{R}(a) = \fint a\,ds^2 \quad \text{is a } \textit{trace} \text{ on } \mathcal{A}\,. \tag{1.550}$$

We then have the following result.

LEMMA 1.151. *Assuming* (1.549) (1.550), *one has*

$$\fint [D^2,a][D^2,b]\,ds^6 = \fint [D,a][D,b]\,ds^4 \qquad \forall\, a,b \in \mathcal{A}\,.$$

PROOF. One has $\mathcal{R}(ab) = \mathcal{R}(b\,a)$ by hypothesis. Moreover, using the trace property of $\fint$, one has

$$\mathcal{R}(b\,a) = \fint b\,a\,D^{-2} = \fint a\,D^{-2}\,b\,D^2\,D^{-2}\,. \tag{1.551}$$

Thus, we get

$$\fint a\,D^{-2}\,[D^2,b]\,D^{-2} = 0 \qquad \forall\, a,b \in \mathcal{A}\,. \tag{1.552}$$

By Lemma 1.150 one has

$$\fint D^{-2}\,a\,[D^2,b]\,D^{-2} = \fint a[D^2,b]\,ds^4 = -\fint [D,a][D,b]\,ds^4.$$

Substracting this from (1.552) we get

$$\fint [D,a][D,b]\,ds^4 = \fint [a,D^{-2}][D^2,b]\,D^{-2}. \tag{1.553}$$

On the other hand we have

$$[a,D^{-2}] = D^{-2}\,[D^2,a]\,D^{-2}\,. \tag{1.554}$$

Now $D^{-2}\,[D^2,a]\,D^{-2}\,[D^2,b]\,D^{-2}$ is of order -4 and we can permute the D^{-2} freely so that we get the statement. □

COROLLARY 1.152. *Assuming* (1.549) *and* (1.550), *for* $a^0, a^1, a^2 \in \mathcal{A}$, *one has*

$$\int\!\!\!\!\!- a^0\,[D,a^1][D,a^2]\,ds^4 = \int\!\!\!\!\!- a^0\,[D^2,a^1][D^2,a^2]\,ds^6\,.$$

PROOF. Let

$$\varphi(a_0,a_1) = \int\!\!\!\!\!-[D,a_0][D,a_1]\,ds^4$$
$$\psi(a_0,a_1) = \int\!\!\!\!\!-[D^2,a_0][D^2,a_1]\,ds^6.$$

We compute the Hochschild coboundary of φ and ψ as

$$\begin{aligned} &\varphi(a_0a_1,a_2) - \varphi(a_0,a_1a_2) + \varphi(a_2a_0,a_1) \\ = \;& \int\!\!\!\!\!-\{(a_0\,[D,a_1][D,a_2] + [D,a_0]\,a_1\,[D,a_2]) - ([D,a_0]\,a_1\,[D,a_2] \\ + \;& [D,a_0][D,a_1]\,a_2) + (a_2[D,a_0][D,a_1] + [D,a_2]\,a_0\,[D,a_1])\}\,ds^4. \end{aligned}$$

Since we can permute freely with ds^4 we get

$$(b\varphi)(a_0,a_1,a_2) = 2\int\!\!\!\!\!- a_0\,[D,a_1][D,a_2]\,ds^4\,. \tag{1.555}$$

The same argument gives

$$(b\psi)(a_0,a_1,a_2) = 2\int\!\!\!\!\!- a_0\,[D^2,a_1][D^2,a_2]\,ds^6\,. \tag{1.556}$$

By Lemma 1.151 one has $\varphi = \psi$ and the conclusion follows. ☐

We can now show the following result on the dependence of the term

$$\int\!\!\!\!\!-(D+A)^{-2}$$

on the inner fluctuations.

PROPOSITION 1.153. *Assuming* (1.549) *and* (1.550), *one has the following properties.*

- *The term* $\int\!\!\!\!- (D+A)^{-2}$ *is independent of* A.
- *The functional* $\mathcal{R}$ *is unchanged by the inner fluctuation* $D \mapsto D + A$.

PROOF. By (1.545) and (1.546) one has

$$\begin{aligned} &\int\!\!\!\!\!- (D+A)^{-2} - \int\!\!\!\!\!- D^2 \\ &= -\int\!\!\!\!\!- (DA+AD)\,D^{-4} + \int\!\!\!\!\!- (DA+AD)^2\,D^{-6} - \int\!\!\!\!\!- A^2\,D^{-4}\,. \end{aligned} \tag{1.557}$$

The first term in the right-hand side vanishes by (1.549).
Now, for $A = \sum_i a_i\,[D,b_i]$, one has

$$DA + AD = \sum_i a_i\,[D^2,b_i] + \sum_i [D,a_i][D,b_i], \tag{1.558}$$

where the second term is of lower order. It follows then from Corollary 1.152 that

$$\int\!\!\!\!\!- (DA_1 + A_1D)(DA_2 + A_2D)\,ds^6 = \int\!\!\!\!\!- A_1\,A_2\,ds^4, \tag{1.559}$$

for any $A_1, A_2 \in \Omega_D^1$. Indeed, it is enough to check this on $A_1 = a_1\,[D, b_1]$ and $A_2 = a_2\,[D, b_2]$. By (1.558), the left-hand side gives

$$\begin{aligned}
⨍ a_1\,[D^2, b_1]\,a_2\,[D^2, b_2]\,ds^6 &= \frac{1}{2}\,b\psi(a_1, b_1 a_2, b_2) - \frac{1}{2}\,b\psi(a_1 b_1, a_2, b_2) \\
&= \frac{1}{2}\,b\varphi(a_1, b_1 a_2, b_2) - \frac{1}{2}\,b\varphi(a_1 b_1, a_2, b_2) \\
&= ⨍ a_1\,[D, b_1]\,a_2\,[D, b_2]\,ds^4.
\end{aligned}$$

which is indeed equal to the right-hand side.
Taking $A_1 = A_2 = A$ we see that the last two terms of (1.557) add up to zero and we get the first statement.
For the second statement, using the invariance of $⨍$ one gets

$$\begin{aligned}
⨍ x(D+A)^{-2} - ⨍ x\,D^{-2} &= -⨍ x\,D^{-2}(DA + AD)\,D^{-2} \\
&+ \int x(DA+AD)^2\,D^{-6} \\
&- ⨍ x\,A^2\,D^{-4} \qquad \forall\, x \in \mathcal{A}\,. \qquad (1.560)
\end{aligned}$$

Now in (1.559) we can replace A_1 by xA_1 to get

$$⨍ x\,(DA_1 + A_1 D)(DA_2 + A_2 D)\,ds^6 = ⨍ x\,A_1\,A_2\,ds^4\,. \qquad (1.561)$$

This shows that the last two terms of (1.560) cancel. To prove the vanishing of the first term of the right-hand side in (1.560), we can assume that $A = a\,[D, b]$. One has $DA + AD = a\,[D^2, b] + [D, a][D, b]$ and

$$⨍ x\,D^{-2}(DA + AD)\,D^{-2} = ⨍ x\,D^{-2}\,a\,[D^2, b]\,D^{-2} + ⨍ x\,D^{-2}\,[D, a][D, b]\,D^{-2}\,.$$

Now, by Corollary 1.152, the last term is

$$⨍ x\,[D, a][D, b]\,D^{-4} = ⨍ x\,[D^2, a][D^2, b]\,D^{-6}\,.$$

One has

$$\begin{aligned}
⨍ x\,D^{-2}\,a\,[D^2, b]\,D^{-2} &= ⨍ x\,a\,D^{-2}\,[D^2, b]\,D^{-2} \\
&- ⨍ x\,D^{-2}\,[D^2, a]\,D^{-2}\,[D^2, b]\,D^{-2}\,.
\end{aligned}$$

The first term in the right-hand side vanishes by (1.550), while the second is

$$-⨍ x\,[D^2, a][D^2, b]\,ds^6$$

and cancels the above. $\square$

We have thus shown that the scalar curvature is insensitive to inner fluctuation of the metric. Notice, however, that this property depends crucially on the assumption of the tadpole vanishing (1.549).

11.2. Seeley–DeWitt coefficients and Gilkey's theorem.

Let $(\mathcal{A}, \mathcal{H}, D)$ be a noncommutative geometry described by a spectral triple. We apply to the square D^2 of the Dirac operator the standard local formulae for the heat expansion (see [**144**] §4.8). We begin by recalling briefly the statement of Gilkey's Theorem ([**144**] Theorem 4.8.16).

One starts with a compact Riemannian manifold M of dimension m, with metric g and one lets V be a vector bundle on M and P a differential operator acting on sections of V and with leading symbol given by the metric tensor. Thus, locally one has

$$P = -\left(g^{\mu\nu} I\, \partial_\mu \partial_\nu + A^\mu \partial_\mu + B\right), \tag{1.562}$$

where $g^{\mu\nu}$ plays the role of the inverse metric, I is the identity matrix, and A^μ and B are endomorphisms of the bundle V.

The Seeley-DeWitt coefficients are the terms $a_n(x, P)$ in the heat expansion, which is of the form

$$\mathrm{Tr}\, e^{-tP} \sim \sum_{n \geq 0} t^{\frac{n-m}{2}} \int_M a_n(x, P)\, dv(x), \tag{1.563}$$

where m is the dimension of the manifold and $dv(x) = \sqrt{\det g_{\mu\nu}(x)}\, d^m x$, with $g_{\mu\nu}$ the metric on M.

By Lemma 4.8.1 of [**144**], the operator P is uniquely written in the form

$$P = \nabla^* \nabla - E \tag{1.564}$$

where ∇ is a connection on V, with $\nabla^* \nabla$ the connection Laplacian, and E is an endomorphism of V.

Let $\Gamma^\rho_{\mu\nu}$ be the Christoffel symbols of the Levi–Civita connection of the metric g. We set

$$\Gamma^\rho := g^{\mu\nu} \Gamma^\rho_{\mu\nu}. \tag{1.565}$$

The explicit formulae for the connection ∇ and the endomorphism E are then of the form

$$\nabla_\mu = \partial_\mu + \omega'_\mu \tag{1.566}$$

$$\omega'_\mu = \frac{1}{2}\, g_{\mu\nu} (A^\nu + \Gamma^\nu \cdot \mathrm{id}), \tag{1.567}$$

with id the identity endomorphism of V, and

$$E = B - g^{\mu\nu} (\partial_\mu\, \omega'_\nu + \omega'_\mu\, \omega'_\nu - \Gamma^\rho_{\mu\nu}\, \omega'_\rho). \tag{1.568}$$

One lets Ω be the curvature of the connection ∇ so that (cf. [**144**] Lemma 4.8.1)

$$\Omega_{\mu\nu} = \partial_\mu\, \omega'_\nu - \partial_\nu\, \omega'_\mu + [\omega'_\mu, \omega'_\nu]. \tag{1.569}$$

We denote multiple covariant differentiation as in [**144**] Lemma 2.4.3, by

$$E_{;\mu}{}^{\mu} := \nabla_\mu \nabla^\mu E. \tag{1.570}$$

The Seeley-DeWitt coefficients $a_n(P)$ vanish for odd values of n. The first three a_n's for n even have the following explicit form in terms of the Riemann curvature tensor R, the curvature Ω of the connection ∇ and the endomorphism E.

THEOREM 1.154. [**144**] *One has :*

$$a_0(x,P) = (4\pi)^{-m/2}\mathrm{Tr}(\mathrm{id}) \tag{1.571}$$

$$a_2(x,P) = (4\pi)^{-m/2}\mathrm{Tr}\left(-\frac{R}{6}\,\mathrm{id}+E\right) \tag{1.572}$$

$$\begin{aligned} a_4(x,P) &= (4\pi)^{-m/2}\frac{1}{360}\mathrm{Tr}(-12R_{;\mu}{}^{\mu}+5R^2-2R_{\mu\nu}\,R^{\mu\nu} \\ &+ 2R_{\mu\nu\rho\sigma}\,R^{\mu\nu\rho\sigma}-60\,R\,E+180\,E^2+60\,E_{;\mu}{}^{\mu} \\ &+ 30\,\Omega_{\mu\nu}\,\Omega^{\mu\nu}). \end{aligned} \tag{1.573}$$

REMARK 1.155. *Notice that in Theorem 1.154 the endomorphism E only appears through the expressions*

$$\mathrm{Tr}\left(-\frac{R}{6}\,\mathrm{id}+E\right) \quad \textit{and} \quad \mathrm{Tr}\left((-\frac{R}{6}\,\mathrm{id}+E)^2\right) \tag{1.574}$$

and the boundary term $\mathrm{Tr}(E_{;\mu}{}^{\mu})$.

11.3. The generalized Lichnerowicz formula.

Let M be a compact Riemannian spin manifold of dimension m, S the spinor bundle with the canonical Riemannian connection ∇_S. Let V be a Hermitian vector bundle over M with a compatible connection ∇_V. One lets $\partial\!\!\!/_V$ be the Dirac operator on $S\otimes V$ endowed with the tensor product connection ([**203**] Proposition 5.10)

$$\nabla(\xi\otimes v) = (\nabla_S\,\xi)\otimes v+\,\xi\otimes(\nabla_V\,v). \tag{1.575}$$

Let then R_V be the bundle endomorphism of the bundle $S\otimes V$ defined by

$$R_V(\xi\otimes v) = \frac{1}{2}\sum_{j,\,k=1}^{m}(\gamma_j\gamma_k\,\xi)\otimes(R(V)_{jk}\,v) \tag{1.576}$$

where $R(V)$ is the curvature tensor of the bundle V.

One then has the following result ([**203**] Theorem 8.17).

THEOREM 1.156. *let s be the scalar curvature of M as in* (1.541). *Then the Dirac operator $\partial\!\!\!/_V$ satisfies*

$$\partial\!\!\!/_V^2 = \nabla^*\,\nabla+\frac{1}{4}\,s+\,R_V, \tag{1.577}$$

where $\nabla^*\,\nabla$ is the connection Laplacian of $S\otimes V$.

Notice that all three terms on the right-hand side of (1.577) are self-adjoint operators by construction. In particular R_V is self-adjoint. One can write R_V in the form

$$R_V = \sum_{j<k}\gamma_j\gamma_k\otimes R(V)_{jk}, \tag{1.578}$$

where the terms in the sum are pairwise orthogonal for the natural inner product on the Clifford algebra induced by the Hilbert-Schmidt inner product $\langle A,B\rangle = \mathrm{Tr}(A^*B)$ in the spin representation. Notice that here $R(V)$ is antisymmetric and the γ_j, γ_k anticommute.

11.4. The Einstein–Yang–Mills system.

In this section we describe a result of [**46**], §2, which illustrates the properties and the role of the spectral action functional in the simplest possible noncommutative modification of an ordinary 4-dimensional manifold, namely the noncommutative geometry obtained by replacing the algebra $C^\infty(M)$ of smooth functions on M by the tensor product

$$\mathcal{A} = C^\infty(M) \otimes M_N(\mathbb{C}), \tag{1.579}$$

with $M_N(\mathbb{C})$ the algebra of complex $N \times N$ matrices. The spectral triple is obtained by tensoring the Dirac spectral triple $(C^\infty(M), L^2(M,S), \partial\!\!\!/_M)$ with the spectral triple for $M_N(\mathbb{C})$ given by the left action of $M_N(\mathbb{C})$ on the Hilbert space of $N \times N$ matrices endowed with the Hilbert-Schmidt norm, and with the trivial Dirac operator $D = 0$. The real structure is given by the adjoint operation $T \mapsto T^*$ on matrices, while J_M denotes the real structure on $(C^\infty(M), L^2(M,S), \partial\!\!\!/_M)$ given by the charge conjugation operator.
Thus, the product geometry is of the form

- $\mathcal{A} = C^\infty(M, M_N(\mathbb{C}))$
- $\mathcal{H} = L^2(M,S) \otimes M_N(\mathbb{C})$
- $J(\xi \otimes T) = J_M \xi \otimes T^*$
- $D = \partial\!\!\!/_M \otimes 1$
- $\gamma = \gamma_5 \otimes 1$.

We compare the spectral action functional with the action for an SU(N) Yang-Mills theory coupled to fermions in the adjoint representation and to Einstein gravity.

We begin by computing the inner fluctuations of the metric for the product geometry $(\mathcal{A}, \mathcal{H}, D)$ described above.
The Dirac operator $\partial\!\!\!/_M$ is of the form

$$\partial\!\!\!/_M = \sqrt{-1}\,\gamma^\mu\,\nabla^s_\mu. \tag{1.580}$$

Here ∇^s is the spin connection which we express in a vierbein e. Namely, the vierbein e^a_μ is defined by the relation

$$g_{\mu\nu} = e^a_\mu e^b_\nu \delta_{ab}, \tag{1.581}$$

so that $\gamma_a = e^\mu_a \gamma_\mu$ satisfy the anticommutation $\{\gamma_a, \gamma_b\} = 2\delta_{ab}$. One sets

$$\gamma_{ab} := \frac{1}{4}[\gamma_a, \gamma_b]. \tag{1.582}$$

They satisfy

$$[\gamma_{ab}, \gamma_{cd}] = \delta_{cb}\gamma_{ad} - \delta_{ca}\gamma_{bd} - \delta_{db}\gamma_{ac} + \delta_{da}\gamma_{bc}$$

and define the spinor representation of the Lie algebra $\mathfrak{o}(4)$. The spin connection is then of the form

$$\nabla^s_\mu = \partial_\mu + \frac{1}{2}\omega^{ab}_\mu \gamma_{ab}. \tag{1.583}$$

The expressions of the form

$$\sum_j a_j\,[D, b_j] \quad \text{with} \quad a_j, b_j \in \mathcal{A}$$

give all operators of the form

$$A = A_\mu\,\gamma^\mu,$$

where the $A_\mu \in C^\infty(M, M_N(\mathbb{C}))$ are arbitrary. The self-adjoint condition $A = A^*$ for gauge potentials is equivalent to $A_\mu = A_\mu^*$, i.e. to $A_\mu \in C^\infty(M, M_N(\mathbb{C}))_{\rm sa}$, the self-adjoint elements of the algebra $\mathcal{A}$. Here the A_μ are acting by left multiplication on $\mathcal{H} = L^2(M, S) \otimes M_N(\mathbb{C})$.

It is convenient for the following to introduce the related notation

$$\mathbb{A}_\mu := -\sqrt{-1}\,\mathrm{ad}\,A_\mu. \tag{1.584}$$

We also denote by ω_μ the spin-connection on M as in (1.583), with

$$\omega_\mu := \frac{1}{2}\,\omega_\mu^{ab}\,\gamma_{ab}. \tag{1.585}$$

Notice that the inner product in $\mathcal{H}$ is given fiberwise using the Hilbert-Schmidt norm on matrices, i.e. by

$$\langle \xi, \eta\rangle = \int_M \mathrm{Tr}(\xi(x)^*\eta(x))\sqrt{g}d^4x,$$

ignoring the spinor index for simplicity. One checks directly that, for $X \in C^\infty(M, M_N(\mathbb{C}))$, the adjoint of left multiplication by X is left multiplication by X^*.

We can then prove the following result on the product geometry

$$(C^\infty(M, M_N(\mathbb{C})), L^2(M, S) \otimes M_N(\mathbb{C}), \partial\!\!\!/_M \otimes 1)\,.$$

PROPOSITION 1.157. *Let $(\mathcal{A}, \mathcal{H}, D)$ be the product geometry described above.*

1. *The inner fluctuations of the metric are parameterized by an* $\mathrm{SU}(N)$ *Yang-Mills field A.*
2. *The operator $D_A = D + A + JAJ^*$ is given by*

$$D_A = \sqrt{-1}\,e_a^\mu\,\gamma^a\left((\partial_\mu + \omega_\mu)\otimes 1_{N^2} + 1 \otimes \mathbb{A}_\mu\right)\,, \tag{1.586}$$

   with $\mathbb{A}_\mu$ as in (1.584).
3. *The coupling of the Yang-Mills field A with the fermions is:*

$$\langle\bar\psi, D_A\psi\rangle \qquad \psi\,,\bar\psi \in \mathcal{H}\,. \tag{1.587}$$

PROOF. (1) We already saw that the gauge potentials are given by $A = A_\mu\gamma^\mu$ with $A_\mu \in C^\infty(M, M_N(\mathbb{C}))_{\rm sa}$ (the self-adjoint elements in the algebra), acting by left multiplication. In dimension 4 one has ([**132**] 5.24)

$$J_M\,\gamma^\mu\,J_M^{-1} = -\gamma^\mu. \tag{1.588}$$

Moreover, the action of J on matrices being given by $T \to T^*$, it transforms the operator of left multiplication by $X \in C^\infty(M, M_N(\mathbb{C}))$ into the right multiplication by X^*. Thus, this shows that

$$A + JAJ^* = \gamma^\mu\,\mathrm{ad}\,A_\mu.$$

Notice that the formula $D_A = D + A + JAJ^*$ eliminates the U(1) part of A, even though one starts with a $U(N)$ gauge potential. Thus, it is natural to impose the "unimodularity" condition

$$\mathrm{Tr}(A) = 0. \tag{1.589}$$

(2) This follows from the same computation above.

(3) We obtained indeed the correct adjoint action on fermions. Notice also that in order to formulate the Euclidean functional integral one considers the $\psi\,,\bar\psi \in \mathcal{H}$ as independent variables [**61**]. ∎

Recall that the Yang–Mills action is of the form

$$S_{YM}(A) = \frac{1}{4}\int_M \mathcal{L}_{YM}(A), \quad \text{with} \quad \mathcal{L}_{YM}(A) = \mathrm{Tr}(F_{\mu\nu}\overline{F}^{\mu\nu}). \tag{1.590}$$

The following result of [**45**] gives the explicit form of the spectral action for the product geometry $(C^\infty(M, M_N(\mathbb{C})), L^2(M,S)\otimes M_N(\mathbb{C}), \partial\!\!\!/_M \otimes 1)$, which recovers the Einstein–Yang–Mills system.

Theorem 1.158. *For the product geometry $(\mathcal{A},\mathcal{H},D)$ described above, the spectral action is given by*

$$\mathrm{Tr}(f(D/\Lambda)) \sim \frac{1}{4\,\pi^2}\int_M \mathcal{L}(g_{\mu\nu}, A)\,\sqrt{g}\,d^4x, \tag{1.591}$$

where $\mathcal{L}(g_{\mu\nu}, A)$ is the Lagrangian of the Einstein–Yang–Mills system given by

$$\begin{aligned}\mathcal{L}(g_{\mu\nu}, A) = {} & 2N^2\,\Lambda^4\,f_4 + \tfrac{N^2}{6}\,\Lambda^2\,f_2\,R + \tfrac{f(0)}{6}\,\mathcal{L}_{YM}(A) \\ & -\tfrac{N^2\,f(0)}{80}\,C_{\mu\nu\rho\sigma}\,C^{\mu\nu\rho\sigma}.\end{aligned} \tag{1.592}$$

modulo topological and boundary terms.

Proof. Using (1.577) and (1.578), we get

$$-\,E \;=\; \frac{1}{4}\,s\otimes \underline{1}_{N^2} + \sum_{\mu<\nu}\gamma^\mu\gamma^\nu\otimes F_{\mu\nu} \tag{1.593}$$

$$\Omega_{\mu\nu} \;=\; \frac{1}{4}\,R^{ab}_{\mu\nu}\,\gamma_{ab}\,\otimes 1_{N^2} + \mathrm{id}_4\otimes F_{\mu\nu}\,, \tag{1.594}$$

where the first term in the formula for $\Omega_{\mu\nu}$ is coming from the curvature of the spin connection, which is expressed in terms of the Riemann tensor R and the γ matrices. Also $F_{\mu\nu}$ is the curvature of the connection $\mathbb{A}_\mu$ obtained from A_μ after passing to the adjoint representation as in (1.584).

We can now use Theorem 1.154 to get the first Seeley–DeWitt coefficients. Neglecting boundary terms, we obtain

$$\begin{aligned}
a_0(P) &= \frac{N^2}{4\pi^2}\int_M \sqrt{g}\,d^4x \\
a_2(P) &= -\frac{N^2}{48\pi^2}\int_M s\,\sqrt{g}\,d^4x = \frac{N^2}{48\pi^2}\int_M R\,\sqrt{g}\,d^4x \\
a_4(P) &= \frac{1}{16\pi^2}\cdot\frac{N^2}{360}\int_M (5R^2 - 8R_{\mu\nu}\,R^{\mu\nu} - 7R_{\mu\nu\rho\sigma}\,R^{\mu\nu\rho\sigma})\,\sqrt{g}\,d^4x \\
&+ \frac{1}{24\pi^2}\int_M \mathrm{Tr}(F_{\mu\nu}\overline{F}^{\mu\nu})\,\sqrt{g}\,d^4x.
\end{aligned}$$

Here $N^2 = \dim M_N(\mathbb{C})$. Notice that here the $a_n(P)$ are as in (1.563). These correspond to the coefficients a_α of (1.527) with $\alpha = (n-m)/2$.

To obtain the coefficient $\frac{1}{24\pi^2}$ of the Yang-Mills action with Lagrangian

$$\mathcal{L}_{\mathrm{YM}}(A) = \mathrm{Tr}(F_{\mu\nu}\overline{F}^{\mu\nu}) \geq 0, \tag{1.595}$$

notice that there are two different contributions coming from Theorem 1.154. The contribution of the term $180E^2$ is obtained from the trace of E^2 which is positive since E is Hermitian by construction. The various terms of (1.593) are pairwise orthogonal in the Hilbert–Schmidt inner product, hence the trace of E^2 is the sum

$$\mathrm{Tr}(E^2) = \frac{N^2}{4} R^2 + 4 \sum_{\mu<\nu} \mathrm{Tr}(F_{\mu\nu} \overline{F}^{\mu\nu}) = \frac{N^2}{4} R^2 + 2\mathrm{Tr}(F_{\mu\nu} \overline{F}^{\mu\nu}), \tag{1.596}$$

where the 4 comes from the dimension of spinors, and the last sum is over all indices $\mu\nu$. The contribution of the term $30\,\Omega_{\mu\nu}\,\Omega^{\mu\nu}$ is obtained from the trace of $\Omega_{\mu\nu}\,\Omega^{\mu\nu}$. Using again the orthogonality in the Hilbert-Schmidt inner product, one sees that the latter is given by

$$\mathrm{Tr}(\Omega_{\mu\nu}\,\Omega_{\mu\nu}) = -\frac{N^2}{2} R^{ab}_{\mu\nu} R^{ab}_{\mu\nu} - 4\mathrm{Tr}(F_{\mu\nu} \overline{F}^{\mu\nu}), \tag{1.597}$$

where the minus sign comes from our notation $F_{\mu\nu}\overline{F}^{\mu\nu} = -F_{\mu\nu}F_{\mu\nu}$ and the 4 comes, as above, from the dimension of the spinor bundle. This shows that the overall coefficient for $\mathrm{Tr}(F_{\mu\nu}\overline{F}^{\mu\nu})$ in a_4 is

$$\frac{1}{16\,\pi^2}\,\frac{1}{360}\,(180 \times 2 - 30 \times 4) = \frac{1}{24\,\pi^2} > 0.$$

In (1.596) and (1.597) notice that we have

$$\frac{1}{4} R^{ab}_{\mu\nu}\gamma_{ab} = \frac{1}{2}\sum_{a<b} R^{ab}_{\mu\nu}\gamma_{ab},$$

and $\mathrm{Tr}(\gamma_{ab}) = 4$ so that we have

$$\frac{4N^2}{4}\sum_{a<b} R^{ab}_{\mu\nu} R^{ab}_{\mu\nu} = \frac{N^2}{2}\sum_{a,b} R^{ab}_{\mu\nu} R^{ab}_{\mu\nu}.$$

where in (1.597), as usual, we do not write explicitly the sum over repeated indices.

The contributions of the Riemann tensor are collected in the form

$$\begin{aligned} \tfrac{1}{16\pi^2}\cdot\tfrac{N^2}{360}\cdot 4\cdot \quad & (5R^2 - 2R_{\mu\nu}\,R^{\mu\nu} + 2R_{\mu\nu\rho\sigma}\,R^{\mu\nu\rho\sigma} \\ & -60R\tfrac{1}{4}R + 180\tfrac{1}{16}R^2 - 30(\tfrac{1}{8}R^{ab}_{\mu\nu}\,R^{ab}_{\mu\nu})) \\ = \tfrac{1}{16\pi^2}\cdot\tfrac{N^2}{360} \quad & (5R^2 - 8R_{\mu\nu}\,R^{\mu\nu} - 7R_{\mu\nu\rho\sigma}\,R^{\mu\nu\rho\sigma}). \end{aligned} \tag{1.598}$$

Notice the factor $\frac{1}{8}$ in $(\frac{1}{8}R^{ab}_{\mu\nu}\,R^{ab}_{\mu\nu})$ in (1.598). This comes from the computation of the trace $\mathrm{Tr}(\Omega_{\mu\nu}\,\Omega_{\mu\nu})$. In fact, one rewrites (1.594) as

$$\Omega_{\mu\nu} = \frac{1}{2}\sum_{a<b} R^{ab}_{\mu\nu}\,\gamma_{ab}\,\otimes 1_{N^2} + \mathrm{id}_4 \otimes F_{\mu\nu},$$

so that the contribution of the Riemann tensor to $\mathrm{Tr}(\Omega_{\mu\nu}\,\Omega_{\mu\nu})$ is of the form

$$-N^2\cdot 4\cdot\frac{1}{4}\sum_{a<b} R^{ab}_{\mu\nu}\,R^{ab}_{\mu\nu} = -N^2\cdot 4\cdot\frac{1}{8}\,R^{ab}_{\mu\nu}\,R^{ab}_{\mu\nu}.$$

Since we are working in dimension four, we have a relation between the Pontrjagin class and the three possible curvature squared terms, of the form

$$R^*R^* := R_{\mu\nu\rho\sigma}\,R^{\mu\nu\rho\sigma} - 4R_{\mu\nu}\,R^{\mu\nu} + R^2.$$

Here the term

$$R^*R^* = \frac{1}{4}\,\varepsilon^{\mu\nu\rho\sigma}\,\varepsilon_{\alpha\beta\gamma\delta}\,R^{\alpha\beta}_{\mu\nu}\,R^{\gamma\delta}_{\rho\sigma} \tag{1.599}$$

is the form that integrates to the Euler characteristic

$$\chi(M) = \frac{1}{32\pi^2} \int R^* R^* \sqrt{g}\, d^4x. \tag{1.600}$$

Moreover, we can change the expression for $a_4(P)$ in terms of $C_{\mu\nu\rho\sigma}$ instead of $R_{\mu\nu\rho\sigma}$, where

$$C_{\mu\nu\rho\sigma} = R_{\mu\nu\rho\sigma} - (g_{\mu[\rho} R_{\nu|\sigma]} - g_{\nu[\rho} R_{\mu|\sigma]}) + \frac{1}{6}(g_{\mu\rho}\, g_{\nu\sigma} - g_{\mu\sigma}\, g_{\nu\rho}) R \tag{1.601}$$

is the Weyl tensor. Using the identity

$$R_{\mu\nu\rho\sigma}\, R^{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma}\, C^{\mu\nu\rho\sigma} + 2R_{\mu\nu}\, R^{\mu\nu} - \frac{1}{3} R^2,$$

we can recast the contribution of the Riemann tensor to $a_4(P)$ into the alternative form

$$\frac{N^2}{16\pi^2} \int (-\frac{1}{20}\, C_{\mu\nu\rho\sigma}\, C^{\mu\nu\rho\sigma} + \frac{11}{360} R^* R^*)\, \sqrt{g}\, d^4x.$$

This ends the proof of (1.592). □

Notice that (1.592) has the correct factor of 2 in the cosmological term because of the definition of f_4, cf. (1.533).

We see in particular that the Yang-Mills terms appear additively in the action (1.592). We explain in §11.5 below that this is in fact a general phenomenon that occurs in the framework of noncommutative geometry.

11.5. Scale independent terms in the spectral action.

In this section we consider the term in Λ^0, i.e. the Λ-independent term in the spectral action and its behavior under inner fluctuations of the metric.

As explained above the gauge potentials make good sense in the framework of noncommutative geometry and come from the inner fluctuations of the metric. This gives meaning to the Feynman graphs all of whose internal lines are fermionic lines, such as the triangle graph of Figure 32.

In the work [**48**] of Chamseddine–Connes, it is shown how the spectral action behaves under the inner fluctuations. We recall here the main results of [**48**] in the following form.

- In the case of a spectral triple $(\mathcal{A}, \mathcal{H}, D)$ with simple dimension spectrum consisting of the positive integers ≤ 4, the variation under inner fluctuations of the Λ^0-term in the spectral action gives the local counterterms for the fermionic graphs

$$\begin{aligned} \zeta_{D+A}(0) - \zeta_D(0) = \; & -⨍ AD^{-1} + \frac{1}{2} ⨍ (AD^{-1})^2 \\ & -\tfrac{1}{3} ⨍ (AD^{-1})^3 + \frac{1}{4} ⨍ (AD^{-1})^4. \end{aligned} \tag{1.602}$$

- Under the hypothesis of vanishing of the Λ^0-tadpole term, i.e. of the tadpole graph of Figure 31, the variation (1.602) is the sum of a Yang–Mills action and a Chern–Simons action relative to a cyclic 3-cocycle on the algebra $\mathcal{A}$.

More precisely, one has the following result (cf. [**48**]).

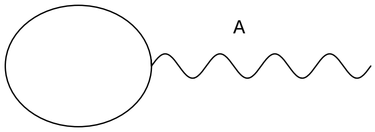


FIGURE 31. The tadpole graph.

THEOREM 1.159. *Let A be a gauge potential for a noncommutative geometry $(\mathcal{A}, \mathcal{H}, D)$ with simple dimension spectrum consisting of positive integers.*

(1) *The function $\zeta_{D+A}(s)$ extends to a meromorphic function with at most simple poles, which is regular at $s = 0$.*
(2) *One has*

$$\zeta_{D+A}(0) - \zeta_D(0) = - ⨍ \log(1 + A\, D^{-1}) = \sum_n \frac{(-1)^n}{n} ⨍ (A\, D^{-1})^n. \tag{1.603}$$

We refer the reader to [**48**] for the detailed computations that give the proof of Theorem 1.159.

REMARK 1.160. In the paper [**47**] the only hypothesis used is that both D and $D+A$ are invertible operators and one can always reduce to this case by adding ϵ to D. Notice that, even in the interpolation $D^2 + tX$ of Lemma 2.3 of [**47**] one does not use anything else, since writing $D^2 + tX = (1-t)D^2 + t(D+A)^2$ shows that one has a uniform lower bound. To define the ζ-function in the presence of zero modes, one can take the limit of the ζ-functions for $D + \epsilon$ for $\epsilon \to 0$. This gives the right limit that counts the zero modes as in Lemma 1.144 above. The main formula of Theorem 1.159 for the variation of zeta in terms of residues remains unaltered.

We now restrict to the case of metric dimension ≤ 4 and we make the *tadpole vanishing* hypothesis (cf. Figure 31)

$$⨍ a\,[D,b]\,D^{-1} = 0\,, \quad \forall\, a\,,\ b \in \mathcal{A}. \tag{1.604}$$

By Theorem 1.159 this condition is equivalent to the vanishing of the first order variation of the Λ-independent part of the spectral action under inner fluctuations, and is therefore a natural hypothesis.
Recall that, by Definition (1.123), the hypothesis of metric dimension ≤ 4 means that

$$D^{-1} \in \mathcal{L}^{(4,\infty)}, \tag{1.605}$$

i.e. that the operator D^{-1} is an infinitesimal of order $\frac{1}{4}$ (cf. [**68**]).
The functional

$$\tau_0(a^0, a^1, a^2, a^3, a^4) = ⨍ a^0\,[D, a^1]\,D^{-1}[D, a^2]\,D^{-1}[D, a^3]\,D^{-1}[D, a^4]\,D^{-1} \tag{1.606}$$

is then a Hochschild cocycle and is given as a Dixmier trace of infinitesimals of order one.

We introduce the notation

$$\alpha(a) = D\,a\,D^{-1} \quad \forall a \in \mathcal{A}. \tag{1.607}$$

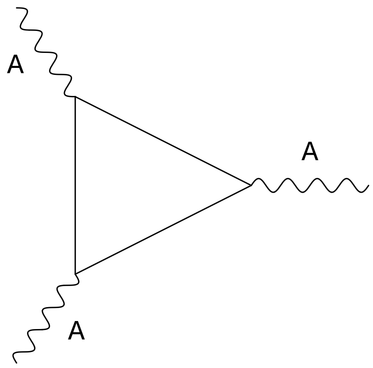


FIGURE 32. The triangle graph.

Notice that in general $\alpha(a) \notin \mathcal{A}$. One has

$$\alpha(ab) = \alpha(a)\,\alpha(b) \qquad a, b \in \mathcal{A}.$$

We then have the following result of [**48**].

THEOREM 1.161. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a spectral triple as in Theorem 1.159. Under the tadpole vanishing hypothesis* (1.604) *the following holds.*

(1) $\psi = \varphi + \frac{1}{2} B_0 \tau_0$ *is a cyclic* 3*-cocycle given by*

$$\psi(a_0, a_1, a_2, a_3) = \frac{1}{2} ⨍ (\alpha(a_0)\, a_1\, \alpha(a_2)\, a_3 - \, a_0\, \alpha(a_1)\, a_2\, \alpha(a_3)), \tag{1.608}$$

with $\alpha(x) = D\, x\, D^{-1}$.

(2) *For any $A \in \Omega_D^1$ one has*

$$-⨍ \log(1 + AD^{-1}) = \frac{1}{4} \int_{\tau_0} (dA + A^2)^2 - \frac{1}{2} \int_{\psi} (A\, dA + \frac{2}{3} A^3). \tag{1.609}$$

Combining this result with Theorem 1.159 one obtains the following result.

COROLLARY 1.162. *In dimension* 4*, the variation under inner fluctuations of the Λ-independent terms of the spectral action is given by*

$$\zeta_{D+A}(0) - \zeta_D(0) = \frac{1}{4} \int_{\tau_0} (dA + A^2)^2 - \frac{1}{2} \int_{\psi} (A\, dA + \frac{2}{3} A^3).$$

This is the sum of the Yang-Mills action described in §*10.6 relative to the Hochschild* 4*-cocycle τ_0 with the Chern-Simons action of* §*10.7, relative to the cyclic* 3*-cocycle ψ.*

Notice that there is still some freedom in the choice of the cocycles τ_0 and ψ involved in Theorem 1.161. Indeed, let $B = A\, B_0$ be the fundamental boundary operator in cyclic cohomology, as discussed in §10.4 above. One has then the following result (cf. [**48**]).

PROPOSITION 1.163. *Let $(\mathcal{A}, \mathcal{H}, D)$ be a geometry as above.*

(1) *Theorem* (1.161) *still holds after replacing $\tau_0 \mapsto \tau_0 + \rho$ and $\psi \mapsto \psi + \frac{1}{2} B_0 \rho$, for any Hochschild* 4*-cocycle ρ such that $B_0 \rho$ is already cyclic, i.e. such that $A\, B_0 \rho = 4\, B_0 \rho$.*

(2) *If the cocycle ψ is in the image of B, i.e. if $\psi \in B(Z^4(\mathcal{A}, \mathcal{A}^*))$, then one can eliminate ψ by a redefinition of τ_0.*

The following lemma shows that "generically" one expects that indeed the cocycle ψ is in the image of B.

LEMMA 1.164. *Let ψ be the cyclic 3-cocycle of Theorem 1.161. Then its pairing with the K_1-group vanishes identically, namely*

$$\langle \psi, u \rangle = 0, \qquad \forall u \in K_1(\mathcal{A})$$

PROOF. The lemma follows directly from Proposition 1.140. $\square$

In particular, the cocycle ψ is in the image of B in the case when $(\mathcal{A}, \mathcal{H}, D)$ is the spectral triple associated to a compact spin Riemannian 4-manifold.

To be more explicit, one defines in general a Hochschild cocycle ρ by

$$\begin{aligned} -6\,\rho(a^0, a^1, a^2, a^3, a^4) \;&=\; ⨍ a^0[D^2, a^1]\,[D^2, a^2]\,[D, a^3]\,[D, a^4]\,D^{-6} \\ &-\; ⨍ a^0[D, a^1]\,[D^2, a^2]\,[D^2, a^3]\,[D, a^4]\,D^{-6} \\ &+\; ⨍ a^0[D, a^1]\,[D, a^2]\,[D^2, a^3]\,[D^2, a^4]\,D^{-6} \\ &-\; ⨍ a^0[D^2, a^1]\,[D, a^2]\,[D, a^3]\,[D^2, a^4]\,D^{-6}\,. \end{aligned}$$

and one checks that $B_0\rho$ is already cyclic.

One then has the following result of [**48**].

PROPOSITION 1.165. *Let $(\mathcal{A}, \mathcal{H}, D)$ be the spectral triple associated to a compact spin Riemannian 4-manifold M. Then $\psi + \frac{1}{2} B_0\rho = 0$ and the Hochschild cocycle $\tau = \tau_0 + \rho$ is positive and given explicitly by*

$$\tau(a^0, a^1, a^2, a^3, a^4) = -\frac{1}{3\,\pi^2} \int_M \langle a^0 da^1 \wedge da^2, da^3 \wedge da^4 \rangle. \tag{1.610}$$

Here we use the inner product on 2-forms associated to the conformal structure of M. One has, equivalently,

$$\int_M \langle \omega, \omega' \rangle = \int_M \omega \wedge \star\omega'.$$

Notice that the minus sign in (1.610) is crucial to ensure the positivity of τ. Indeed, with the notations of §10.6 the adjoint of a 2-form is given by (1.510), so that the skew-symmetry of forms introduces an additional minus sign. In fact, the proof of Proposition 1.165 given in [**48**] gives a direct relation between τ and the positive cocycle of (1.511).

One can then check that Theorem 1.161 provides another computation of the inner fluctuation of the spectral action in the Riemannian case and gives the additive contribution

$$\frac{1}{24\,\pi^2} \int_M \mathrm{Tr}(F_{\mu\nu} \overline{F}^{\mu\nu})\,\sqrt{g}\,d^4x \;\geq 0$$

of Theorem 1.158.

11.6. Spectral action with dilaton.

Another striking property of the scale independent terms in the spectral action is their conformal invariance. Indeed, it is known for instance that the Weyl action is conformally invariant

$$\int C_{\mu\nu\rho\sigma}(G)\, C^{\mu\nu\rho\sigma}(G)\, \sqrt{G}\, d^4x \;=\; \int C_{\mu\nu\rho\sigma}(g)\, C^{\mu\nu\rho\sigma}(g)\, \sqrt{g}\, d^4x\,, \tag{1.611}$$

where G is the rescaled metric (often called the Einstein metric)

$$G^{\mu\nu} = e^{-2\phi} g^{\mu\nu}\,, \quad G_{\mu\nu} = e^{2\phi} g_{\mu\nu} \tag{1.612}$$

and ϕ a dilaton field.

In [**47**] Chamseddine and Connes showed that the arbitrary mass scale Λ in the spectral action can be made dynamical by introducing a scaling dilaton field. We report here briefly on the formulation and properties of the spectral action with dilaton field, following [**47**].

In its simplest form, as we have seen, the spectral action is a counting of the number of eigenvalues λ of D such that $|\lambda| \leq \Lambda$. This is the case when one takes f in (1.526) to be a cutoff function.
The basic idea of [**47**] is to "localize" this counting by replacing

$$N(\Lambda) = \dim\{D^2 \leq \Lambda^2\} \mapsto N(\rho) = \dim\{D^2 \leq \rho^2\}, \tag{1.613}$$

where $\rho = e^{\phi}$, with ϕ the dilaton field, and $\dim\{D^2 \leq \rho^2\}$ is the number of negative modes of the operator $D^2 - \rho^2 \mathrm{id}$.

One can see the immediate advantage of this localization in the non-compact case, i.e. when the geometric space one is dealing with is non-compact. Dealing with such spaces requires a minor extension of the formalism of spectral triples given in Definition 1.120 above.

Definition 1.166. *A noncompact spectral geometry $(\mathcal{A}, \mathcal{H}, D)$ consists of a non-unital algebra $\mathcal{A}$ acting as bounded operators on a Hilbert space $\mathcal{H}$, with a self-adjoint operator D, such that $[D, a]$ is bounded for all $a \in \mathcal{A}$. The operator D has the property that $a(D - \lambda)^{-1}$ is compact for all $a \in \mathcal{A}$ and $\lambda \notin \mathrm{Spec}(D)$. The sign, real structure, and other properties are defined as in the compact case, after adjoining a unit to the algebra.*

In such non-compact cases, the compact resolvent property of the Dirac operator D now only holds after multiplication by elements of $\mathcal{A}$, thus the spectrum of D can be continuous and one can have $N(\Lambda) = \infty$. It is then natural to modify the definition of the spectral action by choosing the function ρ to be compactly supported so that one has $N(\rho) < \infty$.

The main result of [**47**] is that the spectral action with dilaton is given by the spectral action (without dilaton) applied to the Einstein metric (1.612), with the addition of a kinetic term for the dilaton.

When one applies the same technique to the case of the Standard Model (as analyzed in [**47**]), one finds that in that case one needs to rescale not only the metric (1.612) but also the Higgs field by

$$H = e^{-\phi}\, h\,. \tag{1.614}$$

At the level of the Dirac operator, the transition to the new metric corresponds to the transformation

$$D \mapsto e^{-\phi/2} D e^{-\phi/2}\,. \tag{1.615}$$

The appearance of the kinetic term for the dilaton can be understood in the general framework of Definition 1.147 of the scalar curvature in noncommutative geometry, from the difference between $(e^{-\phi/2}\, D\, e^{-\phi/2})^2$ and $e^{-\phi}\, D^2\, e^{-\phi}$.

Indeed, as in [**47**] Appendix C, the next proposition shows that the shift to the Einstein metric and the appearance of the kinetic term for the dilaton hold quite generally.

PROPOSITION 1.167. *Let* $(\mathcal{A}, \mathcal{H}, D)$ *be a spectral triple, regular with simple dimension spectrum, of metric dimension* 4*. Let* $\rho = e^{\phi}$ *with* $\phi = \phi^*$ *in the center of* $\mathcal{A}$*.*

(1) $(\mathcal{A}, \mathcal{H}, D_\phi = e^{-\phi/2}\, D\, e^{-\phi/2})$ *is a spectral triple and its residue functional is the same as for* $(\mathcal{A}, \mathcal{H}, D)$*.*
(2) *Assuming conditions* (1.549) *and* (1.550) *of §11.1, the scalar curvature functionals* $\mathcal{R}$ *and* $\mathcal{R}_\phi$ *of the two spectral triples are related by*

$$\mathcal{R}_\phi(a) = \mathcal{R}(a\, e^{2\phi}) + \frac{1}{2} ⨍ [D,\, a\, e^{\phi}]\, [D,\, e^{\phi}]\, D^{-4}. \tag{1.616}$$

(3) *With the hypothesis of (2) and with* $\rho = e^{\phi}$*, the coefficient of* $\frac{1}{2}\Lambda^2$ *in the expansion of* $N(\Lambda\, \rho)$ *is the sum of the scalar curvature* $\mathcal{R}_\phi(1)$ *with the (positive) kinetic term*

$$\frac{1}{2} ⨍ |[D,\, \rho]|^2\, D^{-4}.$$

PROOF. 1) The first assertion is straightforward, since $\phi = \phi^*$ is in the center of $\mathcal{A}$. To prove that the residues agree is quite technical and we restrict to the case of ordinary pseudodifferential operators on a manifold where it follows from [**303**].

2) For $a \in \mathcal{A}$, one has

$$⨍ a\, (e^{-\phi/2}\, D\, e^{-\phi/2})^{-2} = ⨍ a\, e^{\phi}\, D^{-1}\, e^{\phi}\, D^{-1} \tag{1.617}$$

and

$$D^{-1}\, e^{\phi} = e^{\phi}\, D^{-1} - D^{-1}\, [D,\, e^{\phi}]\, D^{-1}. \tag{1.618}$$

This makes it possible to write (1.617) in the form

$$⨍ a\, (e^{-\phi/2}\, D\, e^{-\phi/2})^{-2} = ⨍ a\, e^{2\phi}\, D^{-2} - ⨍ a\, e^{\phi}\, D^{-1}\, [D,\, e^{\phi}]\, D^{-2}. \tag{1.619}$$

Then, using (1.618) again, one obtains

$$\begin{aligned} -⨍ a e^{\phi} D^{-1} [D,\, e^{\phi}] D^{-2} = \; & -⨍ a e^{\phi} [D, e^{\phi}] D^{-3} \\ & -⨍ D^{-1} [D, a e^{\phi}] D^{-1} [D, e^{\phi}] D^{-2}. \end{aligned} \tag{1.620}$$

The first of the two terms vanishes by the tadpole condition (1.549). Using $D[D,x] + [D,x]D = [D^2,x]$ and $[D,x]D^{-1} = -D^{-1}[D,x] + D^{-1}[D^2,x]D^{-1}$ for $x = a\, e^{\phi}$, we can write the term

$$T = -\int\!\!\!\!\!\!- a\, e^{\phi}\, [D,\, e^{\phi}]\, D^{-3} \;-\; \int\!\!\!\!\!\!- D^{-1}\, [D,\, a\, e^{\phi}]\, D^{-1}\, [D,\, e^{\phi}]\, D^{-2}$$

on the right-hand side of (1.620) in the form

$$T = \int\!\!\!\!\!\!- \; [D,\, a\, e^{\phi}]\, [D,\, e^{\phi}]\, D^{-4} \;-\; \int\!\!\!\!\!\!- \; D^{-2}\, [D^2,\, a\, e^{\phi}]\, D^{-1}\, [D,\, e^{\phi}]\, D^{-2}. \tag{1.621}$$

Consider now the expression

$$S = \int\!\!\!\!\!\!- \; D^{-2}\, [D^2,\, a\, e^{\phi}]\, D^{-1}\, [D,\, e^{\phi}]\, D^{-2}.$$

We show that

$$S = \frac{1}{2} \int\!\!\!\!\!\!- \; [D^2,\, a\, e^{\phi}]\, [D^2,\, e^{\phi}]\, D^{-6}. \tag{1.622}$$

To see this, notice that the commutator of D with $[D^2,\, a\, e^{\phi}]$ is equal to $[D^2,\, [D,\, a\, e^{\phi}]]$ and has order 1 so that, moving D to the left and using the trace property of the residue, one gets

$$S = \frac{1}{2} \int\!\!\!\!\!\!- \; [D^2,\, a\, e^{\phi}]\, D^{-2}\, (D\, [D,\, e^{\phi}] + [D,\, e^{\phi}]\, D)\, D^{-4}$$

and one obtains (1.622). Summarizing, we have obtained the equality

$$\begin{aligned} \mathcal{R}_{\phi}(a) = \; & \int\!\!\!\!\!\!- a e^{2\phi} D^{-2} + \int\!\!\!\!\!\!- [D,\, a\, e^{\phi}]\, [D,\, e^{\phi}]\, D^{-4} \\ & -\tfrac{1}{2} \int\!\!\!\!\!\!- [D^2,\, a\, e^{\phi}]\, [D^2,\, e^{\phi}]\, D^{-6}. \end{aligned} \tag{1.623}$$

Thus, by Lemma 1.151 one gets (1.616).

3) The condition $D^2 \le \Lambda^2 \rho^2$ means that $e^{-\phi} D^2 e^{-\phi} \le \Lambda^2$. Notice that one cannot use a strict cutoff function such as the characteristic function of an interval and it should be replaced by an approximation by a smooth cutoff. We ignore this subtlety. We are interested in comparing the coefficients of Λ^2 in the heat expansions with $t = \Lambda^{-2}$ of $e^{-\phi} D^2 e^{-\phi}$ and D_{ϕ}^2. This amounts to comparing the residues of their inverses. By 2) we have

$$\int\!\!\!\!\!\!- (e^{-\phi} D^2 e^{-\phi})^{-1} = \mathcal{R}_{\phi}(1) - \frac{1}{2} \int\!\!\!\!\!\!- [D, \rho]^2\, D^{-4}.$$

Since the operator $[D, \rho]$ is anti-Hermitian we then get 3). ☐

12. Noncommutative geometry and the Standard Model

It is well known that it is possible to infer the Minkowski space geometry underlying special relativity from experimental evidence associated to Maxwell's theory of electromagnetism. In fact, the Maxwell equations are intrinsically relativistic. This geometry is then extended to the Lorentzian manifolds of general relativity. From the particle physics viewpoint, the Lagrangian of electromagnetism is just a very small part of the full Standard Model Lagrangian, as one can see easily from

the explicit expression given in §9.4 and §9.6.2 above. Thus, it is natural to wonder whether the transition from the Lagrangian of electrodynamics to the Standard Model can be understood as a further refinement of the geometry of spacetime.
We approach this question using the tools we introduced above, namely

- Noncommutative geometry
- The spectral action principle

As we have seen, noncommutative geometry makes it possible to consider spaces which are more general than ordinary manifolds. This gives us more freedom to obtain a suitable geometric setting that accounts for the additional terms in the Lagrangian.
It is worth pointing out again, as we mentioned in §9.9, the difference between the approach we intend to follow and the usual geometric formulation in terms of bundles. The usual formulation maintains an a priori distinction between gravity and the other forces, while we seek a geometric formulation for the full theory of the Standard Model minimally coupled to gravity and with neutrino mixing which reduces it to *pure gravity* on a "refined spacetime". This in particular means that the symmetries will be the (analog of the) diffeomorphism group on the resulting space and not bundle homomorphisms that are a priori required to preserve the fibered structure.

We work in Euclidean rather than Lorentz signature, leaving as an important problem the Wick rotation back to the Minkowski signature. For a formulation in Minkowski signature see [**13**].

The idea of interpreting the Lagrangian of the Standard Model in terms of noncommutative geometry goes back a good twenty years (Connes [**68**], [**70**], see also [**85**], [**181**], [**180**]). The origin of this interpretation of the Standard Model in terms of noncommutative geometry lies in the extension of the Yang-Mills functional to the algebraic framework of noncommutative geometry (cf. [**68**], [**96**]). The color degrees of freedom were still added somewhat artificially in the early models (cf. [**68**]) and the action functional was obtained by analogy with the classical gauge theories. In particular it did not incorporate gravity. The work [**70**] already shows most of the main points we are going to discuss below, such as the fact that the fermions of the Standard Model provide the Hilbert space $\mathcal{H}$ of a spectral triple for a suitable algebra $\mathcal{A}$, while the bosons arise naturally as inner fluctuations of the corresponding Dirac operator D. As we are going to see in the following sections, the spectral geometry $(\mathcal{A}, \mathcal{H}, D)$ is, in these models, a product of an ordinary geometry by a "finite" noncommutative space $(\mathcal{A}_F, \mathcal{H}_F, D_F)$. Besides a $\mathbb{Z}/2$ grading γ this spectral triple has a crucial piece of structure: a real structure as in Definition 1.124.

The work of Chamseddine–Connes in [**45**], [**46**] and the results of [**70**] show how to incorporate the color naturally and, more importantly, how to obtain the bosonic part of the Standard Model action coupled to Einstein gravity from the very general *spectral* action principle discussed in §11. In fact, the main result of [**45**], [**46**] is that, when applied to the inner fluctuations of a product geometry of the form $M \times F$ the spectral action gives the Standard Model coupled to gravity in the Euclidean form. As explained above the spectral action principle gives a simple prescription to define an action functional and applies in full generality, so that the action can be evaluated on all noncommutative geometries. It is in essence a *purely gravitational* action. Since it only depends upon the spectrum of the line element i.e. the inverse of the

Dirac operator. In physics terms, modulo a chiral gauge transformation, the Dirac operator is the inverse of the Euclidean propagator of fermions. The spectral action principle then asserts that D is all that is needed to define the bosonic part of the action.
Moreover, since disjoint unions of spaces correspond to direct sums of the Dirac operators, a simple additivity requirement for the action functional shows that it has to be of the form

$$S = \mathrm{Tr}(f(D/\Lambda)), \tag{1.624}$$

where f is an even function of the real variable and Λ a parameter fixing the mass scale. In fact, as we saw in §11, the choice of the test function f plays only a small role. In fact, when the action S is expanded in inverse powers of Λ, it depends on the first momenta $\int f(u)\, u^{k-1}\, du$ and on the Taylor expansion of f at 0.
The spectral action functional admits a huge invariance group, namely the group $\mathcal{U}$ of unitary operators on the Hilbert space $\mathcal{H}$ which preserve the additional structures given by the $\mathbb{Z}/2$-grading γ and the real structure J i.e. such that

$$U\,U^* = U^*\,U = 1\,, \quad U\,J = J\,U\,, \quad U\,\gamma = \gamma\,U \tag{1.625}$$

That the gravitational Einstein action appears naturally in the expansion of S is reminiscent of the idea of induced gravity, but the spectral action in fact recovers the full bosonic sector.

We have already seen in §11.4 that the product geometry of a Riemannian spin 4-manifold by the simplest noncommutative geometry

$$(M_N(\mathbb{C}), M_N(\mathbb{C}), 0)$$

gives the $\mathrm{SU}(N)$ Yang-Mills gauge theory minimally coupled to gravity, with fermions in the adjoint representation. The inner fluctuations (taking the real structure J into account) give the gauge bosons and the spectral action gives the required Lagrangian, with the correct signs required for the Euclidean functional integral.
The theory obtained in this way is, however, massless and the key missing step is the Brout–Englert–Higgs mechanism of spontaneous symmetry breaking, which generates masses without spoiling the renormalizability of the theory. We refer to [**60**] for the detailed description of the Higgs mechanism and the specific role of the Goldstone bosons. The lack of a simple geometric interpretation of the new terms added to the Lagrangian to generate the masses leads to many theoretical speculations denying any fundamental significance to the Higgs particle. As we see below, noncommutative geometry indeed provides a simple geometric interpretation of the new terms. The geometry is again a product geometry of a Riemannian spin four-manifold by a finite geometry F but the latter will be more subtle than the case of $(M_N(\mathbb{C}), M_N(\mathbb{C}), 0)$.

For the noncommutative geometry F used in [**46**] to obtain the Standard Model coupled to gravity, all the ingredients are *finite-dimensional*. The algebra $\mathcal{A}_F = \mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C})$ (i.e. the direct sum of the algebras $\mathbb{C}$ of complex numbers, $\mathbb{H}$ of quaternions, and $M_3(\mathbb{C})$ of 3×3 matrices) encodes the gauge group. The Hilbert space $\mathcal{H}_F$ is of dimension 90 and encodes the elementary quarks and leptons. The operator D_F encodes those free parameters of the Standard Model related to the Yukawa couplings.

For M the spectral triple is given by the representation of the algebra of smooth functions acting by multiplication in the Hilbert space $L^2(M,S)$ of square integrable spinors, the grading is given by γ_5 and the real structure J_M is given by charge conjugation.

While it did achieve the purpose of deriving the Standard Model action from simple geometric principles, the work [**46**], [**70**] was incomplete in the following ways.

(1) The finite geometry F was put in "by hand", with no conceptual understanding of the representation of $\mathcal{A}_F$ in $\mathcal{H}_F$ nor of the possible choices of the operator D_F.
(2) There is a fermion doubling problem (cf. [**210**]) in the fermionic part of the action.
(3) This formulation does not incorporate the neutrino mixing and see-saw mechanism for neutrino masses.

We present here a new formulation, based on the recent work [**73**], [**52**], which provides a solution for these three problems (the first only partly, but the full solution ([**49**], [**50**]) will be explained briefly in §18.3). In terms of noncommutative geometry, the main new point is that, when dealing with noncommutative spaces one should keep the distinction between the following two notions of dimension.

- The metric dimension
- The KO-dimension.

The metric dimension manifests itself through the dimension spectrum and its upper bound as discussed in §10. In the model we consider below, the metric dimension of the product geometry is 4, the same as the ordinary space-time manifold, i.e. the metric dimension of the finite noncommutative geometry F is zero, as is guaranteed by the finite dimensionality of both $\mathcal{A}_F$ and $\mathcal{H}_F$.

The KO-dimension is only well defined modulo 8 and it takes into account both the $\mathbb{Z}/2$-grading γ of $\mathcal{H}$ as well as the real structure J (cf. §10.8). The real surprise is that in order for things to work we have to set up the real structure of the finite geometry F so that its KO-dimension is equal to 6. It is only thanks to this that the fermion doubling problem pointed out in [**210**] can be successfully handled. Moreover this will give rise to a geometry that automatically generates through the spectral action the full Standard Model minimally coupled to gravity, including neutrino mixing and the see-saw mechanism.

As a result, we see that the KO-dimension of the product space $M \times F$ is in fact equal to $10 \sim 2$ modulo 8. In particular, looking back at the table of Definition 1.124, we get that the antilinear isometry J fulfills $J^2 = -1$. Thus, the above unitary group $\mathcal{U}$ of (1.625) is now related to the symplectic group preserving the quaternionic structure associated to J, where the quaternion j in the traditional basis of $\mathbb{H}$ over $\mathbb{R}$, given by (i,j,k), acts in $\mathcal{H}$ by the operator J.

The fact of having a space of dimension 10 is reminiscent of the models of string theory, with the finite noncommutative geometry F as a replacement for the Calabi-Yau fibers that are at the core of string theory. This 6-dimensional noncommutative geometry can perhaps be realized as a low energy truncation of the type of compact fibers of string theory models, as considered for instance in [**135**]. On the other hand, the KO-dimension 10 is also 2 modulo 8, which may be related to the observations of [**202**] about gravity.

13. The finite noncommutative geometry

Our approach to the Standard Model is based on the idea of providing *only* a very simple mathematical input, a fairly natural choice of a finite-dimensional algebra, and then *deriving* from that input alone the full complexity of the Standard Model Lagrangian coupled to gravity by applying the formalism of noncommutative geometry and the spectral action.

Input	Machinery	Output
algebra $\mathcal{A}_{LR}$	noncommutative geometry spectral triples spectral action principle	Standard Model with neutrino mixing coupled to gravity

Our initial input is what we refer to as the "left–right symmetric algebra". Notice that the model we are going to derive to account for massive neutrinos and neutrino oscillations is not a left–right symmetric model, as pointed out already in §9.6 above. In fact, as we see below in §13.1, a mathematical mechanism will *select* a subalgebra of the left–right symmetric algebra, which breaks the left–right symmetry and provides the actual algebra $\mathcal{A}_F$ that will appear in the spectral triple.
We shall explain later in §18.3 how the classification of the finite noncommutative geometries of KO-dimension 6 leads more conceptually to the same result.

Definition 1.168. *The left-right symmetric algebra $\mathcal{A}_{LR}$ is the involutive finite-dimensional algebra*

$$\mathcal{A}_{LR} := \mathbb{C} \oplus \mathbb{H}_L \oplus \mathbb{H}_R \oplus M_3(\mathbb{C}). \tag{1.626}$$

Here $\mathbb{H}_L$ and $\mathbb{H}_R$ are two copies of the real algebra $\mathbb{H}$ of quaternions, with the indices L, R for bookkeeping. The involution is given by

$$(\lambda, q_L, q_R, m)^* := (\bar{\lambda}, \bar{q}_L, \bar{q}_R, m^*), \tag{1.627}$$

where $q \mapsto \bar{q}$ denotes the involution on the algebra of quaternions. The complex subalgebra $\mathbb{C} \oplus M_3(\mathbb{C})$ is referred to as the integer spin part and the real subalgebra $\mathbb{H}_L \oplus \mathbb{H}_R$ as the half-integer spin part of $\mathcal{A}_{LR}$.

We recall for later use that the algebra $\mathbb{H}$ of quaternions has a basis $\{1, i\sigma^\alpha\}$, where σ^α are the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.628}$$

It is important to stress here that the main advantage of working with *associative algebras* as opposed to the more common use of *Lie algebras* in physics is that the representation theory is much more restrictive. In particular, a finite-dimensional algebra has only a finite number of irreducible representations, hence a canonical representation in their sum. Thus, having only specified the datum of the algebra, in our case the left-right symmetric algebra of Definition 1.168, one obtains *canonically* an associated bimodule, which we now describe in the case of interest to us.
We begin by introducing the adjoint action of an algebra $\mathcal{A}$ on a $\mathcal{A}$-bimodule $\mathcal{M}$.

Definition 1.169. *Let $\mathcal{M}$ be a bimodule over an involutive algebra $\mathcal{A}$. For $u \in \mathcal{A}$ unitary, i.e. such that $uu^* = u^*u = 1$, one defines* $\mathrm{Ad}(u)$ *by* $\mathrm{Ad}(u)\xi = u\xi u^*$, *for all $\xi \in \mathcal{M}$.*

In the specific case of the left-right symmetric algebra $\mathcal{A}_{LR}$, we also introduce the notion of odd bimodule. This corresponds to "half-integer spin" representations.

DEFINITION 1.170. *Let $\mathcal{M}$ be an $\mathcal{A}_{LR}$-bimodule. Then $\mathcal{M}$ is* odd *iff the adjoint action (as in Definition 1.169) of the element $s = (1, -1, -1, 1)$ fulfills* $\mathrm{Ad}(s) = -1$.

We give a description of odd bimodules in terms of a complex algebra obtained from $\mathcal{A}_{LR}$ and its opposite algebra.

LEMMA 1.171. *Let $\mathcal{A}^0_{LR}$ denote the opposite algebra of $\mathcal{A}_{LR}$ and let $\mathcal{B} = (\mathcal{A}_{LR} \otimes_{\mathbb{R}} \mathcal{A}^0_{LR})_p$ denote the reduction of $\mathcal{A}_{LR} \otimes_{\mathbb{R}} \mathcal{A}^0_{LR}$ by the projection*

$$p = \frac{1}{2}\,(1 - s \otimes s^0).$$

Then $\mathcal{B}$ is an algebra over $\mathbb{C}$ isomorphic to the direct sum of four copies of the algebra $M_2(\mathbb{C}) \oplus M_6(\mathbb{C})$.

PROOF. By construction, the algebra $\mathcal{B}$ is of the form

$$\mathcal{B} = (\mathbb{H}_L \oplus \mathbb{H}_R) \otimes_{\mathbb{R}} (\mathbb{C} \oplus M_3(\mathbb{C}))^0 \oplus (\mathbb{C} \oplus M_3(\mathbb{C})) \otimes_{\mathbb{R}} (\mathbb{H}_L \oplus \mathbb{H}_R)^0. \tag{1.629}$$

Thus, the result follows from the isomorphisms

$$\mathbb{H} \otimes_{\mathbb{R}} \mathbb{C} = M_2(\mathbb{C}) \quad \text{and} \quad \mathbb{H} \otimes_{\mathbb{R}} M_3(\mathbb{C}) = M_6(\mathbb{C}). \tag{1.630}$$

□

The algebras $M_N(\mathbb{C})$ and $\mathbb{H}$ are isomorphic to their opposite algebras, respectively by $m \mapsto m^t$ for matrices and $q \mapsto \bar{q}$ for quaternions. We use this anti-isomorphism to obtain a representation π^0 of the opposite algebra $\mathcal{A}^0_{LR}$ from a representation π of $\mathcal{A}_{LR}$. We then have the following description of odd bimodules.

LEMMA 1.172. *Let $\mathcal{B} = (\mathcal{A}_{LR} \otimes_{\mathbb{R}} \mathcal{A}^0_{LR})_p$ be as in Lemma 1.171. Then, any odd bimodule $\mathcal{M}$ for $\mathcal{A}_{LR}$ is a representation of $\mathcal{B}$.*

PROOF. The statement follows directly from the action of

$$s = (1, -1, -1, 1)$$

in Definition 1.170. □

We adopt here the physicists's convention and we denote an irreducible representation by its dimension in boldface.

DEFINITION 1.173. *Let $\mathbf{3}$ denote the 3-dimensional irreducible representation of the algebra $M_3(\mathbb{C})$ and $\mathbf{3}^0$ that of the opposite algebra. Let $\mathbf{1}$ and $\mathbf{1}^0$ denote respectively the 1-dimensional irreducible representation of $\mathbb{C}$ and its opposite, and $\mathbf{2}$ and $\mathbf{2}^0$ the 2-dimensional irreducible representation of the algebra $\mathbb{H}$ and of its opposite. One writes $\mathbf{2}_L$ and $\mathbf{2}_R$ to distinguish between $\mathbb{H}_L$ and $\mathbb{H}_R$.*

Since $\mathcal{B} = (\mathcal{A}_{LR} \otimes_{\mathbb{R}} \mathcal{A}^0_{LR})_p$ is an algebra over $\mathbb{C}$, we restrict ourselves to considering those odd bimodules $\mathcal{M}$ that are complex representations of $\mathcal{B}$. We then introduce the notion of *contragredient* bimodule.

DEFINITION 1.174. *The contragredient bimodule $\mathcal{M}^0$ of an $\mathcal{A}_{LR}$-bimodule $\mathcal{M}$ is the complex conjugate space*

$$\mathcal{M}^0 = \{\bar{\xi}\ ;\ \xi \in \mathcal{M}\}\,, \quad a\,\bar{\xi}\,b = \overline{b^*\xi\,a^*}\,, \quad \forall\, a,\ b \in \mathcal{A}_{LR}. \tag{1.631}$$

As explained above, it is now natural to consider the representation of the algebra $\mathcal{A}_{LR}$ on the bimodule $\mathcal{M}_F$, obtained as the direct sum of all inequivalent irreducible odd $\mathcal{A}_{LR}$-bimodules. Notice that the reason for working with bimodules instead of just left or right modules is that we are going to construct a spectral triple endowed with a real structure and the condition (1.471) related to the involution J of the real structure by (1.472) signals the presence of a bimodule structure. We describe more explicitly the bimodule $\mathcal{M}_F$.

PROPOSITION 1.175. *Let $\mathcal{M}_F$ be the direct sum of all inequivalent irreducible odd $\mathcal{A}_{LR}$-bimodules.*

- *The dimension of the complex vector space $\mathcal{M}_F$ is* 32.
- *The $\mathcal{A}_{LR}$-bimodule $\mathcal{M}_F = \mathcal{E} \oplus \mathcal{E}^0$ is the direct sum of the bimodule*

$$\mathcal{E} := \mathbf{2}_L \otimes \mathbf{1}^0 \oplus \mathbf{2}_R \otimes \mathbf{1}^0 \oplus \mathbf{2}_L \otimes \mathbf{3}^0 \oplus \mathbf{2}_R \otimes \mathbf{3}^0 \tag{1.632}$$

  with its contragredient $\mathcal{E}^0$.
- *The $\mathcal{A}_{LR}$-bimodule $\mathcal{M}_F$ is isomorphic to the contragredient bimodule $\mathcal{M}_F^0$ by the antilinear isometry J_F given by*

$$J_F(\xi, \bar{\eta}) = (\eta, \bar{\xi}) , \quad \forall \xi \,,\ \eta \in \mathcal{E} \tag{1.633}$$

- *One has*

$$J_F^2 = 1 \,, \quad \xi\, b = J_F b^* J_F\, \xi \,, \quad \forall \xi \in \mathcal{M}_F \,,\ b \in \mathcal{A}_{LR} \tag{1.634}$$

PROOF. The first two statements follow from the structure of the algebra $\mathcal{B}$ described in Lemma 1.171. In fact, the complex algebra $M_N(\mathbb{C})$ admits only one irreducible representation and the latter has dimension N. Thus, the sum of the irreducible representations of $\mathcal{B}$ is given by

$$\mathbf{2}_L \otimes \mathbf{1}^0 \oplus \mathbf{2}_R \otimes \mathbf{1}^0 \oplus \mathbf{2}_L \otimes \mathbf{3}^0 \oplus \mathbf{2}_R \otimes \mathbf{3}^0 \oplus \mathbf{1} \otimes \mathbf{2}_L^0 \oplus \mathbf{1} \otimes \mathbf{2}_R^0 \oplus \mathbf{3} \otimes \mathbf{2}_L^0 \oplus \mathbf{3} \otimes \mathbf{2}_R^0, \tag{1.635}$$

of dimension $4 \times 2 + 4 \times 6 = 32$. Notice then that, by construction, $\mathcal{M}_F$ is the direct sum $\mathcal{E} \oplus \mathcal{E}^0$ of the bimodule (1.632) with its contragredient, and that the map (1.633) gives the required antilinear isometry. Notice, moreover, that (1.634) follows from (1.631). □

In addition to the anti-involution J_F of (1.633) we have a natural grading given as follows.

COROLLARY 1.176. *Let γ_F be the $\mathbb{Z}/2$-grading on $\mathcal{M}_F$ defined by*

$$\gamma_F = c - J_F\, c\, J_F \,, \quad \text{with} \quad c = (0, 1, -1, 0) \in \mathcal{A}_{LR}, \tag{1.636}$$

where J_F is as in (1.633). *The pair (J_F, γ_F) satisfies*

$$J_F^2 = 1 \quad \text{and} \quad J_F\, \gamma_F = -\, \gamma_F\, J_F. \tag{1.637}$$

In particular, the relations (1.637) *imply that the KO-dimension of the finite geometry is equal to* 6 *mod* 8.

PROOF. This follows directly from the definition of J_F in (1.633). The table of Definition 1.124 shows that the signs in the relations (1.637) can only happen in the case of KO-dimension 6 modulo 8. □

The fact that the algebra $\mathcal{A}_{LR}$ is a natural choice of input for our model can be justified as follows. The algebras $\mathbb{H}_L \oplus \mathbb{H}_R$ and $M_2(\mathbb{C})$ are closely related since they have the same complexification. Thus the algebra $\mathcal{A}_{LR}$ of (1.626) is strikingly close to the truncation $\oplus_{n=1}^{3} M_n(\mathbb{C})$ of the group ring of SU(2). In particular, the element of the group ring corresponding to the kernel of the covering $\mathrm{SU}(2) \to SO(3)$ is the element $s = (1, -1, -1, 1) \in \mathcal{A}_{LR}$ that we considered in Definition 1.170, where the action of s in an irreducible representation is ± 1 according to the parity of the spin, consistently with our notion of *odd bimodule* in Definition 1.170. The relation to the group ring of SU(2) suggests possible connections with the theory of quantum groups at roots of unity (cf. e.g. [**54**] §11 and §9.3.C). We shall give in §18.3 a completely different conceptual explanation for taking as an input, instead of $\mathcal{A}_{LR}$, the algebra $M_2(\mathbb{H}) \oplus M_4(\mathbb{C})$ which leads, by the order one condition of §13.1, to exactly the same finite noncommutative geometry F.

In the following section we construct a finite spectral triple, starting from the representation of $\mathcal{A}_{LR}$ on a bimodule which is a direct sum

$$\mathcal{H}_F = \mathcal{M}_F^{\oplus N} \tag{1.638}$$

of N copies of $\mathcal{M}_F$. The number N will be identified with the number of particle generations in the Standard Model, and we take in fact $N = 3$, so that $\mathcal{H}_F$ has dimension 96. At this point the choice of $N = 3$ is an extra input, which is not deduced from our initial input $\mathcal{A}_{LR}$.

13.1. The subalgebra and the order one condition.

As we explain in §13.2 below, the two terms in the decomposition

$$\mathcal{M}_F = \mathcal{E} \oplus \mathcal{E}^0 \tag{1.639}$$

of Proposition 1.175 correspond in the physical interpretation to particles and antiparticles for the fermions of the Standard Model. One can see from the explicit form of the representation that the action of the algebra $\mathcal{A}_{LR}$ preserves the splitting (1.639) and J simply interchanges $\mathcal{E}$ with its contragredient $\mathcal{E}^0$. Thus, the construction so far never mixes the sectors $\mathcal{E}$ and $\mathcal{E}^0$. The important part of the geometry that contains some nontrivial intertwining of these spaces will be encoded in the Dirac operator of the finite spectral triple. The very natural requirement that the Dirac operator indeed has nontrivial "off-diagonal pieces" that intertwine the $\mathcal{E}$ and $\mathcal{E}^0$ will select a maximal subalgebra of $\mathcal{A}_{LR}$.

To see this, we consider as in (1.638) the module $\mathcal{H}_F = \mathcal{M}_F^{\oplus 3}$. This inherits induced J_F and γ_F operators from those on $\mathcal{M}_F$ defined by (1.633) and (1.636).

Definition 1.177. *We set*

$$\mathcal{H}_f := \mathcal{E} \oplus \mathcal{E} \oplus \mathcal{E}, \quad \mathcal{H}_{\bar{f}} := \mathcal{E}^0 \oplus \mathcal{E}^0 \oplus \mathcal{E}^0, \tag{1.640}$$

so that we have

$$\mathcal{H}_F = \mathcal{H}_f \oplus \mathcal{H}_{\bar{f}}. \tag{1.641}$$

The following simple observation shows that, indeed, the action of $\mathcal{A}_{LR}$ preserves the matter/antimatter splitting of $\mathcal{H}_F$ induced by (1.639).

LEMMA 1.178. *The left action of $\mathcal{A}_{LR}$ splits as the sum of a representation π on $\mathcal{H}_f$ and a representation π' on $\mathcal{H}_{\bar{f}}$.*
These representations of $\mathcal{A}_{LR}$ are disjoint*, i.e. they have no equivalent subrepresentations.*

PROOF. This follows directly from the explicit description of the bimodule $\mathcal{M}_F$ as in Proposition 1.175. ☐

It is important, in the context of noncommutative geometry and spectral triples, to think of the data $\mathcal{A}$ and D of the algebra and the Dirac operator as a *coupled system*, which one solves for pairs of $\mathcal{A}$ and D satisfying certain natural properties. In our setting, as we mentioned above, we want the Dirac operator to intertwine the spaces $\mathcal{H}_f$ and $\mathcal{H}_{\bar{f}}$ while satisfying the compatibility with the algebra, i.e. the order one condition (1.473), which is the only compatibility condition in the finite-dimensional case, the bounded commutators condition being automatically satisfied. The next Lemma gives a description of the conditions on the pair of $\mathcal{A} \subset \mathcal{A}_{LR}$ and D that fulfill these requirements. For any operator $T : \mathcal{H}_f \to \mathcal{H}_{\bar{f}}$ we let

$$\mathcal{A}(T) = \{b \in \mathcal{A}_{LR} \mid \pi'(b)T = T\pi(b)\,,\ \pi'(b^*)T = T\pi(b^*)\}. \tag{1.642}$$

It is by construction an involutive unital subalgebra of $\mathcal{A}_{LR}$.

LEMMA 1.179. *Let $\mathcal{A} \subset \mathcal{A}_{LR}$ be an involutive unital subalgebra of $\mathcal{A}_{LR}$. Then the following properties hold.*

1. *If the restrictions of π and π' to $\mathcal{A}$ are disjoint, then there is no off-diagonal Dirac operator for $\mathcal{A}$.*
2. *If there exists an off-diagonal Dirac operator for $\mathcal{A}$, then there exists a pair e, e' of minimal projections in the commutants of $\pi(\mathcal{A}_{LR})$ and $\pi'(\mathcal{A}_{LR})$ and an operator T such that $e'Te = T \neq 0$ and $\mathcal{A} \subset \mathcal{A}(T)$.*

PROOF. 1) First the order one condition shows that $[D, a^0]$ cannot have an off-diagonal part since it is in the commutant of $\mathcal{A}$. Conjugating by J shows that $[D, a]$ cannot have an off-diagonal part. Thus the off-diagonal part D_{off} of D commutes with $\mathcal{A}$ i.e. $[D_{\text{off}}, a] = 0$, and $D_{\text{off}} = 0$ since there are no intertwining operators.
2) By 1) the restrictions of π and π' to $\mathcal{A}$ are not disjoint and there exists a non-zero operator T such that $\mathcal{A} \subset \mathcal{A}(T)$. For any elements x, x' of the commutants of π and π', one has

$$\mathcal{A}(T) \subset \mathcal{A}(x'Tx)$$

since $\pi'(b)T = T\pi(b)$ implies $\pi'(b)x'Tx = x'Tx\pi(b)$. Taking a partition of unity by minimal projections there exists a pair e, e' of minimal projections in the commutants of π and π' such that $e'Te \neq 0$ so that one can assume $e'Te = T \neq 0$. ☐

In particular, Lemma 1.179 shows that fixing $\mathcal{A} = \mathcal{A}_{LR}$ precludes the existence of operators D in $\mathcal{H}_F$ that fulfill the order one condition (1.473) and intertwine the subspaces $\mathcal{H}_f$ and $\mathcal{H}_{\bar{f}}$.
The existence of such operators D intertwining $\mathcal{H}_f$ and $\mathcal{H}_{\bar{f}}$ is restored by passing to a unique subalgebra of maximal dimension in $\mathcal{A}_{LR}$, as the following result shows.

PROPOSITION 1.180. *Up to an automorphism of $\mathcal{A}_{LR}$, there exists a unique involutive subalgebra $\mathcal{A}_F \subset \mathcal{A}_{LR}$ of maximal dimension admitting off-diagonal Dirac* *operators, namely operators that intertwine the subspaces $\mathcal{H}_f$ and $\mathcal{H}_{\bar{f}}$ of $\mathcal{H}_F$. The subalgebra is given by*

$$\mathcal{A}_F = \{(\lambda, q_L, \lambda, m) \mid \lambda \in \mathbb{C}\,,\ q_L \in \mathbb{H}\,,\ m \in M_3(\mathbb{C})\} \sim \mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C}). \tag{1.643}$$

PROOF. Let $\mathcal{A} \subset \mathcal{A}_{LR}$ be an involutive unital subalgebra. If it admits an off-diagonal Dirac, then by Lemma 1.179 it is contained in a subalgebra $\mathcal{A}(T)$ with the support of T contained in a minimal projection of the commutant of $\pi(\mathcal{A}_{LR})$ and the range of T contained in the range of a minimal projection of the commutant of $\pi'(\mathcal{A}_{LR})$.
This reduces the argument to two cases, where the representation π is the irreducible representation of $\mathbb{H}$ on $\mathbb{C}^2$ and π' is either the representation of $\mathbb{C}$ on $\mathbb{C}$ or the irreducible representation of $M_3(\mathbb{C})$ on $\mathbb{C}^3$.
In the first case the support E of T is one-dimensional. The commutation relation (1.642) defines the subalgebra $\mathcal{A}(T)$ from the condition $\lambda T\xi = Tq\xi$, for all $\xi \in E$, which implies $\lambda\xi - q\xi = 0$. Thus, in this case the algebra $\mathcal{A}(T)$ is the pullback of

$$\{(\lambda, q) \in \mathbb{C} \oplus \mathbb{H} \mid q\,\xi = \lambda\,\xi\,, \quad \forall \xi \in E\} \tag{1.644}$$

under the projection on $\mathbb{C} \oplus \mathbb{H}$ from $\mathcal{A}_{LR}$. The algebra (1.644) is the graph of an embedding of $\mathbb{C}$ in $\mathbb{H}$. Such an embedding is unique up to inner automorphisms of $\mathbb{H}$. In fact, the embedding is determined by the image of $i \in \mathbb{C}$ and all elements in $\mathbb{H}$ satisfying $x^2 = -1$ are conjugate.
The corresponding subalgebra $\mathcal{A}_F \subset \mathcal{A}_{LR}$ is of real codimension 4. Up to the exchange of the two copies of $\mathbb{H}$ it is given by (1.643).
In the second case the operator T has at most two-dimensional range $\mathcal{R}(T)$. This range is invariant under the action π' of the subalgebra $\mathcal{A}$ and so is its orthogonal complement since $\mathcal{A}$ is involutive. Thus, in all cases the $M_3(\mathbb{C})$-part of the subalgebra is contained in the algebra of $2 \oplus 1$ block diagonal 3×3 matrices, which is of real codimension 8 in $M_3(\mathbb{C})$. Hence $\mathcal{A}$ is of codimension at least $8 > 4$ in $\mathcal{A}_{LR}$.
The actual existence of Dirac operators for the subalgebra (1.643) with off-diagonal terms follows from Theorem 1.187 below, which gives the complete classification of all Dirac operators compatible with the data

$$(\mathcal{A}_F, \mathcal{H}_F, J_F, \gamma_F).$$

$\square$

13.2. The bimodule $\mathcal{H}_F$ and fermions.

We introduce a more explicit notation for the elements of the bimodule $\mathcal{H}_F$, which we can then relate explicitly to the fermions of the Standard Model through the assignment of hypercharges in §13.3 below.

As before, we denote by **2** the 2-dimensional irreducible representation of the algebra $\mathbb{H}$ of quaternions of the form

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \tag{1.645}$$

with $\alpha, \beta \in \mathbb{C}$.

DEFINITION 1.181. *Let $|\uparrow\rangle$ and $|\downarrow\rangle$ be the basis of the irreducible representation* $\mathbf{2}$ *of* $\mathbb{H}$ *of (1.645) for which the action of $\lambda \in \mathbb{C} \subset \mathbb{H}$ is diagonal with eigenvalues λ on $|\uparrow\rangle$ and $\bar{\lambda}$ on $|\downarrow\rangle$. Let $u = (u_j)$, and $\bar{u} = (\bar{u}_j)$, with $j = 1, 2, 3$ the color index, denote the elements*

$$u \in |\uparrow\rangle \otimes \mathbf{3}^0 \subset \mathbf{2} \otimes \mathbf{3}^0 \quad \text{and} \quad \bar{u} \in \mathbf{3} \otimes |\uparrow\rangle \subset \mathbf{3} \otimes \mathbf{2}^0 \tag{1.646}$$

One writes $|\uparrow\downarrow\rangle_L$ or $|\uparrow\downarrow\rangle_R$ to distinguish the $\mathbb{H}_L$ and $\mathbb{H}_R$ cases and correspondingly writes u_L and u_R for the elements in $|\uparrow\rangle_L \otimes \mathbf{3}^0$ and $|\uparrow\rangle_R \otimes \mathbf{3}^0$ in $\mathcal{E}$ and d_L and d_R for the elements in $|\downarrow\rangle_L \otimes \mathbf{3}^0$ and $|\downarrow\rangle_R \otimes \mathbf{3}^0$ in $\mathcal{E}$. Similarly for $\bar{u}_L$, $\bar{u}_R$, $\bar{d}_L$, and $\bar{d}_R$ in $\mathcal{E}^0$. For $\lambda = 1, \ldots, N$ the generation index, one lets u_L^λ, u_R^λ, d_L^λ, d_R^λ, $\bar{u}_L^\lambda$, $\bar{u}_R^\lambda$, $\bar{d}_L^\lambda$, $\bar{d}_R^\lambda$ be the corresponding elements in $(\mathcal{E} \oplus \mathcal{E}^0) \otimes \mathbb{C}^N$.
Let then ν and e denote the elements

$$\nu \in |\uparrow\rangle \otimes \mathbf{1}^0 \subset \mathbf{2} \otimes \mathbf{1}^0 \quad \text{and} \quad \bar{\nu} \in \mathbf{1} \otimes |\uparrow\rangle \subset \mathbf{1} \otimes \mathbf{2}^0, \tag{1.647}$$

$$e \in |\downarrow\rangle \otimes \mathbf{1}^0 \subset \mathbf{2} \otimes \mathbf{1}^0 \quad \text{and} \quad \bar{\nu} \in \mathbf{1} \otimes |\downarrow\rangle \subset \mathbf{1} \otimes \mathbf{2}^0. \tag{1.648}$$

As above one writes ν_L, ν_R, $\bar{\nu}_L$, $\bar{\nu}_R$, and e_L, e_R, $\bar{e}_L$, $\bar{e}_R$, to distinguish between $\mathbb{H}_L$ and $\mathbb{H}_R$, with an additional generation index $\lambda = 1, \ldots, N$, so that one has e_L^λ, ν_L^λ, e_R^λ, ν_R^λ, $\bar{e}_L^\lambda$, $\bar{\nu}_L^\lambda$, $\bar{e}_R^\lambda$, $\bar{\nu}_R^\lambda$ in $(\mathcal{E} \oplus \mathcal{E}^0) \otimes \mathbb{C}^N$.

Upon writing the bimodule $\mathcal{M}$ as $\mathcal{M} = \mathcal{E} \oplus \mathcal{E}^0$, with $\mathcal{E}$ as in (1.632), and setting $N = 3$, we obtain in this way a description of elements of $\mathcal{H}_F$ as combinations of elements of the form

$$\begin{array}{llll} u_L^\lambda, & d_L^\lambda, & u_R^\lambda, & d_R^\lambda, \\ \bar{u}_L^\lambda, & \bar{d}_L^\lambda, & \bar{u}_R^\lambda, & \bar{d}_R^\lambda, \\ \nu_L^\lambda, & e_L^\lambda, & \nu_R^\lambda, & e_R^\lambda, \\ \bar{\nu}_L^\lambda, & \bar{e}_L^\lambda, & \bar{\nu}_R^\lambda, & \bar{e}_R^\lambda. \end{array} \tag{1.649}$$

Physically, the u_L^λ, d_L^λ, u_R^λ, d_R^λ, $\bar{u}_L^\lambda$, $\bar{d}_L^\lambda$, $\bar{u}_R^\lambda$, $\bar{d}_R^\lambda$ describe particles and antiparticles in the *quark sector*

$$\big(\mathbf{2}_L \otimes \mathbf{3}^0 \oplus \mathbf{2}_R \otimes \mathbf{3}^0 \oplus \mathbf{3} \otimes \mathbf{2}_L^0 \oplus \mathbf{3} \otimes \mathbf{2}_R^0\big) \otimes \mathbb{C}^3 \tag{1.650}$$

with both possible chiralities. The e_L^λ, ν_L^λ, e_R^λ, ν_R^λ, $\bar{e}_L^\lambda$, $\bar{\nu}_L^\lambda$, $\bar{e}_R^\lambda$, $\bar{\nu}_R^\lambda$ describe particles and antiparticles in the *lepton sector*

$$\big(\mathbf{2}_L \otimes \mathbf{1}^0 \oplus \mathbf{2}_R \otimes \mathbf{1}^0 \oplus \mathbf{1} \otimes \mathbf{2}_L^0 \oplus \mathbf{1} \otimes \mathbf{2}_R^0\big) \otimes \mathbb{C}^3, \tag{1.651}$$

where we allow for right-handed neutrinos ν_R^λ as well as their antiparticles, unlike the minimal Standard Model, which only allows for one chirality. Notice that, in the lepton sector, the neutrinos ν_L^λ, ν_R^λ, $\bar{\nu}_L^\lambda$, $\bar{\nu}_R^\lambda$ play the role of the "up particles" and the charged leptons e_L^λ, e_R^λ, $\bar{e}_L^\lambda$, $\bar{e}_R^\lambda$ of the "down particles" in the quark sector. This gives a description of the elements of the bimodule $\mathcal{H}_F$ in terms of the fermions of the Standard Model (with right-handed neutrinos) in the following form.

	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$
u_L^λ, u_R^λ	up quark	charm quark	top quark
$\bar{u}_L^\lambda$, $\bar{u}_R^\lambda$	up anti-quark	charm anti-quark	top anti-quark
d_L^λ, d_R^λ	down quark	strange quark	bottom quark
$\bar{d}_L^\lambda$, $\bar{d}_R^\lambda$	down anti-quark	strange anti-quark	bottom anti-quark
ν_L^λ, ν_R^λ	electron neutrino	muon neutrino	τ neutrino
$\bar{\nu}_L^\lambda$, $\bar{\nu}_R^\lambda$	electron anti-neutrino	muon anti-neutrino	τ anti-neutrino
e_L^λ, e_R^λ	electron	muon	τ
$\bar{e}_L^\lambda$, $\bar{e}_R^\lambda$	positron	anti-muon	anti-τ

Notice that, in choosing the basis of fermions there is an ambiguity on whether one multiplies by the mixing matrix for the down particles. This point will be discussed more explicitly in §17 below, see (1.795).

We can then describe more explicitly the action of $\mathcal{A}_F$ on $\mathcal{H}_F$ in terms of the fermions. We use the notation f_L, $\bar{f}_L$, f_R, $\bar{f}_R$ for any of the elements of the form (1.649) and we say that f or $\bar{f}$ is a quark if it is as in the first two lines of (1.649) and that it is a lepton if it is as in the last two lines of (1.649).
Let $a = (\lambda, q, m) \in \mathcal{A}_F$, with $\lambda \in \mathbb{C}$, $q = \alpha + \beta j \in \mathbb{H}$ and $m \in M_3(\mathbb{C})$. We let q^t denote the transpose of the matrix (1.645) of the representation **2** of $\mathbb{H}$. We also denote by $q(\lambda)$ the embedding $\mathbb{C} \subset \mathbb{H}$ given by $\alpha = \lambda$ and $\beta = 0$ and we denote by $q(\lambda)^t = q(\lambda)$ the corresponding (diagonal) matrix as in (1.645). Then we have the following description of the action of $\mathcal{A}_F$ on $\mathcal{H}_F$.

LEMMA 1.182. *The action of $a \in \mathcal{A}_F$ on $\mathcal{H}_f$ is given by*

$$
\begin{aligned}
a\begin{pmatrix} u_L \\ d_L \end{pmatrix} &= q^t \begin{pmatrix} u_L \\ d_L \end{pmatrix} &&= \begin{pmatrix} \alpha u_L - \bar{\beta} d_L \\ \beta u_L + \bar{\alpha}\, d_L \end{pmatrix}, \\
a\begin{pmatrix} u_R \\ d_R \end{pmatrix} &= q(\lambda)^t \begin{pmatrix} u_R \\ d_R \end{pmatrix} &&= \begin{pmatrix} \lambda u_R \\ \bar{\lambda} d_R \end{pmatrix}, \\
a\begin{pmatrix} \nu_L \\ e_L \end{pmatrix} &= q^t \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} &&= \begin{pmatrix} \alpha \nu_L - \bar{\beta} e_L \\ \beta \nu_L + \bar{\alpha}\, e_L \end{pmatrix}, \\
a\begin{pmatrix} \nu_R \\ e_R \end{pmatrix} &= q(\lambda)^t \begin{pmatrix} \nu_R \\ e_R \end{pmatrix} &&= \begin{pmatrix} \lambda \nu_R \\ \bar{\lambda} e_R. \end{pmatrix}
\end{aligned}
\tag{1.652}
$$

The action of $a \in \mathcal{A}_F$ on $\mathcal{H}_{\bar{f}}$ is given by

$$
\begin{aligned}
a\,\bar{f} &= \lambda\,\bar{f} \quad \textit{when } f \textit{ is a lepton,} \\
a\,\bar{f} &= m\,\bar{f} \quad \textit{when } f \textit{ is a quark.}
\end{aligned}
\tag{1.653}
$$

Here the 3×3 matrix m is acting on the color index of quarks.

PROOF. The result follows directly from the explicit description of the bimodule $\mathcal{M}_F$ as in (1.635), after adopting the notation of Definition 1.181. Note that, since we describe the action on basis vectors rather than on coordinates we used the transposed matrices in (1.652). □

Notice how, unlike in the case of the miminal Standard Model, the presence of right-handed neutrinos makes the action on $\mathcal{H}_f$ symmetric in the quark and lepton sectors.

The involution (1.633) of the real structure takes the form

$$J_F(\sum \lambda_i f_i + \sum \mu_j \overline{f}_j) = \sum \overline{\mu}_j f_j + \sum \overline{\lambda}_i \overline{f}_i, \tag{1.654}$$

and the grading (1.636) is given by

$$\gamma_F f_L = f_L, \quad \gamma_F f_R = -f_R, \quad \gamma_F \bar{f}_L = -\bar{f}_L, \quad \gamma_F \bar{f}_R = \bar{f}_R. \tag{1.655}$$

This sign convention for the grading agrees with that of [**285**], while for instance [**233**] follows the opposite sign convention.

13.3. Unimodularity and hypercharges.

We need to justify the identification made in the previous section between the elements of $\mathcal{H}_F$ and the fermions. We do this by showing in this section that this identification gives the correct symmetries of the Standard Model and assigns the correct values of the hypercharges to fermions.

In general one has the following notion.

Definition 1.183. *The unitary group of an involutive algebra $\mathcal{A}$ is given by*

$$U(\mathcal{A}) = \{u \in \mathcal{A} \mid uu^* = u^*u = 1\}. \tag{1.656}$$

In our context, we can define the special unitary group $\mathrm{SU}(\mathcal{A}) \subset U(\mathcal{A})$ as follows.

Definition 1.184. *Let $\mathrm{SU}(\mathcal{A}_F)$ be the subgroup of $U(\mathcal{A}_F)$ defined by*

$$\mathrm{SU}(\mathcal{A}_F) = \{u \in U(\mathcal{A}_F) \mid \det(u) = 1\}, \tag{1.657}$$

where $\det(u)$ is the determinant of the action of u on $\mathcal{H}_F$.

The following result describes the group $\mathrm{SU}(\mathcal{A}_F)$ and its adjoint action.

Proposition 1.185. (1) *The group $\mathrm{SU}(\mathcal{A}_F)$ is of the form*

$$\mathrm{SU}(\mathcal{A}_F) \sim \mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3), \tag{1.658}$$

modulo a finite abelian group.

(2) *The adjoint action of the $\mathrm{U}(1)$ factor is given by multiplication of the basis vectors in $\mathcal{H}_f$ by powers of $\lambda \in \mathrm{U}(1)$, where the powers are given by the following list.*

$$\begin{array}{c|cccc} & |\uparrow\rangle \otimes \mathbf{1}^0 & |\downarrow\rangle \otimes \mathbf{1}^0 & |\uparrow\rangle \otimes \mathbf{3}^0 & |\downarrow\rangle \otimes \mathbf{3}^0 \\ \mathbf{2}_L & -1 & -1 & \frac{1}{3} & \frac{1}{3} \\ \mathbf{2}_R & 0 & -2 & \frac{4}{3} & -\frac{2}{3} \end{array} \tag{1.659}$$

(3) *The adjoint action of the $\mathrm{SU}(2)$ factor is trivial on the right handed particle sector and is the representation $\mathbf{2}$ on the left-handed particle sector.*

(4) *The adjoint action of the $\mathrm{SU}(3)$ factor is trivial on the lepton particle sector and is the color action on the quark particle sector.*

Proof. (1) Let $u = (\lambda, q, m) \in U(\mathcal{A}_F)$. The determinant of the action of u on the subspace $\mathcal{H}_f$ is equal to 1 by construction, since a unitary quaternion has determinant 1. Thus, $\det(u)$ is the determinant of the action $\pi'(u)$ on $\mathcal{H}_{\bar{f}}$. This representation is given by $4 \times 3 = 12$ copies of the irreducible representations $\mathbf{1}$ of

$\mathbb{C}$ and $\mathbf{3}$ of $M_3(\mathbb{C})$, where the 4 is from $\mathbf{2}_L^0 \oplus \mathbf{2}_R^0$ and the 3 is the additional overall multiplicity of the representation given by the number $N = 3$ of generations.
Thus, we have

$$\det(u) = \lambda^{12} \det(m)^{12}.$$

Thus, $\mathrm{SU}(\mathcal{A}_F)$ is the product of the group SU(2), which is the unitary group of $\mathbb{H}$, by the fibered product $G = \mathrm{U}(1) \times_{12} \mathrm{U}(3)$ of pairs $(\lambda, m) \in \mathrm{U}(1) \times \mathrm{U}(3)$, such that $\lambda^{12} \det(m)^{12} = 1$.
One has an exact sequence

$$1 \to \mu_3 \to \mathrm{U}(1) \times \mathrm{SU}(3) \to G \xrightarrow{\mu} \mu_{12} \to 1, \tag{1.660}$$

where μ_N is the group of roots of unity of order N and the maps are as follows. The last map μ is given by $\mu(\lambda, m) = \lambda \det(m)$. By definition of G, the image of the map μ is the group μ_{12} of 12th roots of unity. The kernel of μ is the subgroup $G_0 \subset G$ of pairs $(\lambda, m) \in \mathrm{U}(1) \times \mathrm{U}(3)$ such that $\lambda \det(m) = 1$.
The map $\mathrm{U}(1) \times \mathrm{SU}(3) \to G$ is given by $(\lambda, m) \mapsto (\lambda^3, \lambda^{-1} m)$. Its image is G_0. Its kernel is the subgroup of $\mathrm{U}(1) \times \mathrm{SU}(3)$ of pairs $(\lambda, \lambda\, 1_3)$ where $\lambda \in \mu_3$ is a cube root of 1 and 1_3 is the identity 3×3 matrix.
Thus we obtain an exact sequence of the form

$$1 \to \mu_3 \to \mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3) \to \mathrm{SU}(\mathcal{A}_F) \to \mu_{12} \to 1. \tag{1.661}$$

(2) Up to a finite abelian group, the U(1) factor of $\mathrm{SU}(\mathcal{A}_F)$ is the subgroup of elements of $\mathrm{SU}(\mathcal{A}_F)$ of the form $u(\lambda) = (\lambda, 1, \lambda^{-1/3} 1_3)$, where $\lambda \in \mathbb{C}$, with $|\lambda| = 1$. We ignore the ambiguity in the cube root.
Let us compute the action of $\mathrm{Ad}(u(\lambda))$. One has $\mathrm{Ad}(u) = u(u^*)^0 = ub^0$ with $b = (\bar{\lambda}, 1, \lambda^{1/3} 1_3)$.
This gives the required table as in (1.659) for the restriction to the multiples of the left action $\mathbf{2}_L$. In fact, the left action of u is trivial there.
The right action of $b = (\bar{\lambda}, 1, \lambda^{1/3} 1_3)$ is by $\bar{\lambda}$ on the multiples of $\mathbf{1}^0$ and by $\lambda^{1/3} 1_3^t$ on multiples of $\mathbf{3}^0$.
For the restriction to the multiples of the left action $\mathbf{2}_R$ one needs to take into account the left action of u. This acts by λ on $|\uparrow\rangle$ and $\bar{\lambda}$ on $|\downarrow\rangle$. This adds a ± 1 according to whether the arrow points up or down.
(3) We identify the SU(2) factor with elements of the form $u = (1, q, 1)$ where q is a unitary quaternion. On the particle sector the left action of u is by q on the left-handed part and trivial on the right-handed part. The right action i.e. the action of JuJ^{-1} is trivial. Thus the adjoint action $uJuJ^{-1}$ is $\mathbf{2}$ on left-handed and $\mathbf{1}$ on right handed (for the particle sector).
(4) We identify the SU(3) factor with elements of the form $u = (1, 1, \bar{v})$ where $v \in \mathrm{SU}(3)$ and $\bar{v}$ is the complex conjugate. On the particle sector the left action of u is trivial and the right action JuJ^{-1} is trivial on leptons and is the representation $\mathbf{3}$ on the quark sector. $\square$

Notice how the finite groups μ_3 and μ_{12} in the exact sequence (1.661) are of different nature from the physical viewpoint, the first arising from the center of the color U(3), while the second (μ_{12}) depends upon the presence of three generations.

The result of Proposition 1.185 gives the correct physical assignment of the hypercharges as in (1.415) for all the fermions of the Standard Model, since we have proved that, for $u(\lambda) = (\lambda, 1, \lambda^{-1/3} 1_3)$, we have

$$\mathrm{Ad}(u(\lambda))(e_R) = \bar{\lambda}^2 \, e_R \,, \quad \mathrm{Ad}(u(\lambda))(e_L) = \bar{\lambda} \, e_L \,, \quad \mathrm{Ad}(u(\lambda))(\nu_L) = \bar{\lambda} \, \nu_L$$

which gives the hypercharges Y_ℓ and Y_e of (1.415) and for the quarks one gets the correct values for Y_u and Y_d and Y_q of (1.415) from the action of $\mathrm{Ad}(u) = ub^0$, with $b = (\bar{\lambda}, 1, \lambda^{1/3} 1_3)$, where the b^0 multiplies all the terms by $\lambda^{1/3}$ so that one obtains $Y_u = 1 + 1/3 = 4/3$, $Y_d = -1 + 1/3 = -2/3$, and $Y_q = 0 + 1/3 = 1/3$. The fact that we obtain the correct physical hypercharges justifies our identification of the elements of $\mathcal{H}_F$ with the fermions discussed in §13.2 above.

Thus, the conclusion up to this point is that, through the initial *input* of the algebra $\mathcal{A}_{LR}$ we have *derived* from simple principles the bimodule $\mathcal{M}_F$ (hence $\mathcal{H}_F$ under the assumption of $N = 3$ generations) and the algebra $\mathcal{A}_F$. We have further derived from this setting the symmetries of the Standard Model and obtained the correct fermionic content, with the respective hypercharges. Thus, fermions and hypercharges are a first *output* we obtain from our initial choice of the algebra. We continue now with the remaining part of the structure, which is the Dirac operator of the finite geometry $\mathcal{H}_F$.

13.4. The classification of Dirac operators.

We give a classification of Dirac operators D_F for the pair $(\mathcal{A}_F, \mathcal{H}_F)$. Instead of giving a complete classification, we restrict to those self-adjoint operators D_F on $\mathcal{H}_F$ which, in addition to satisfying the compatibility with $\mathcal{A}_F$ given by the order one condition (1.473), also have the additional property that they commute with the subalgebra $\mathbb{C}_F \subset \mathcal{A}_F$ defined by

$$\mathbb{C}_F = \{(\lambda, \lambda, 0) \,, \lambda \in \mathbb{C}\}. \tag{1.662}$$

The physical reason for imposing this condition is that the commutation relation of the Dirac operator D_F with the algebra $\mathbb{C}_F$ ensures that the photon remains massless. This will be seen more clearly when we discuss how to recover the bosons of the Standard Model from the inner fluctuations of the product geometry in §15.2 and §15.4 below.

We introduce some preliminary notation that will be useful for the classification result in Theorem 1.187 below.

Consider operators on $\mathcal{H}_F$ of the form

$$D(\Upsilon) = \begin{pmatrix} S & T^* \\ T & \bar{S} \end{pmatrix}, \tag{1.663}$$

where the linear map S is written in the form

$$S = S_\ell \oplus (S_q \otimes 1_3), \tag{1.664}$$

with

$$S_\ell : \mathbf{2}_R \otimes \mathbf{1}^0 \oplus \mathbf{2}_L \otimes \mathbf{1}^0 \to \mathbf{2}_R \otimes \mathbf{1}^0 \oplus \mathbf{2}_L \otimes \mathbf{1}^0, \tag{1.665}$$

represented by a matrix of the form

$$S_\ell = \begin{pmatrix} 0 & 0 & \Upsilon_\nu^* & 0 \\ 0 & 0 & 0 & \Upsilon_e^* \\ \Upsilon_\nu & 0 & 0 & 0 \\ 0 & \Upsilon_e & 0 & 0 \end{pmatrix}, \tag{1.666}$$

while

$$S_q \otimes 1_3 : \mathbf{2}_R \otimes \mathbf{3}^0 \oplus \mathbf{2}_L \otimes \mathbf{3}^0 \to \mathbf{2}_R \otimes \mathbf{3}^0 \oplus \mathbf{2}_L \otimes \mathbf{3}^0, \tag{1.667}$$

with S_q represented by a matrix of the form

$$S_q = \begin{pmatrix} 0 & 0 & \Upsilon_u^* & 0 \\ 0 & 0 & 0 & \Upsilon_d^* \\ \Upsilon_u & 0 & 0 & 0 \\ 0 & \Upsilon_d & 0 & 0 \end{pmatrix}. \tag{1.668}$$

Here the Υ_ν, Υ_e, Υ_u, Υ_d and their adjoints are $N \times N$ matrices, with $N = 3$ the number of generations, while 1_3 in (1.664) is the identity 3×3 matrix on the color indices. Similarly, the component T is a linear map defined by setting

$$T : |\uparrow\rangle_R \otimes \mathbf{1}^0 \to \mathbf{1} \otimes |\uparrow\rangle_R, \qquad T(\nu_R) = \Upsilon_R \bar{\nu}_R, \tag{1.669}$$

with Υ_R an $N \times N$ matrix (with $N = 3$ generations) and such that

$$T|_{\mathcal{H}_F \ominus E_R} = 0$$

on the complement of

$$E_R = |\uparrow\rangle_R \otimes \mathbf{1}^0 \subset \mathcal{H}_F. \tag{1.670}$$

In (1.663), T^* denotes the adjoint of T and $\bar{S} = \bar{S}_\ell \oplus (1_3 \otimes \bar{S}_q)$ the corresponding action on $\mathcal{H}_{\bar{f}}$ by the complex conjugate matrices.

We now see that this apparently very special class of operators in fact gives us the general form of Dirac operators for the finite geometry.

In order to proceed to the classification, we first prove the following preliminary result, which determines the commutant $\mathcal{A}'_F$ of $\mathcal{A}_F$ in $\mathcal{H}_F$.

Lemma 1.186. *Consider an operator P on $\mathcal{H}_F = \mathcal{H}_f \oplus \mathcal{H}_{\bar{f}}$, which can always be written in the form*

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}. \tag{1.671}$$

Then P belongs to the commutant, $P \in \mathcal{A}'_F$, iff the following conditions are satisfied.

- *$P_{11} : \mathcal{H}_f \to \mathcal{H}_f$ is block diagonal, with three blocks in $M_{12}(\mathbb{C})$, $M_{12}(\mathbb{C})$, and $1_2 \otimes M_{12}(\mathbb{C})$, corresponding to the subspaces where the action of (λ, q, m) is by $\lambda \in \mathbb{C}$, $\bar{\lambda} \in \mathbb{C}$ and $q \in \mathbb{H}$, respectively.*
- *$P_{12} : \mathcal{H}_{\bar{f}} \to \mathcal{H}_f$ has support in $\mathbf{1} \otimes \mathbf{2}_L^0 \oplus \mathbf{1} \otimes \mathbf{2}_R^0$ and range in $|\uparrow\rangle_R \otimes \mathbf{1}^0 \oplus |\uparrow\rangle_R \otimes \mathbf{3}^0$.*
- *$P_{21} : \mathcal{H}_f \to \mathcal{H}_{\bar{f}}$ has support in $|\uparrow\rangle_R \otimes \mathbf{1}^0 \oplus |\uparrow\rangle_R \otimes \mathbf{3}^0$ and range in $\mathbf{1} \otimes \mathbf{2}_L^0 \oplus \mathbf{1} \otimes \mathbf{2}_R^0$.*
- *$P_{22} : \mathcal{H}_{\bar{f}} \to \mathcal{H}_{\bar{f}}$ is of the form*

$$P_{22} = P_\ell \oplus (1_3 \otimes P_q), \tag{1.672}$$

with

$$P^\ell : \mathbf{1} \otimes (\mathbf{2}_L^0 \oplus \mathbf{2}_R^0) \to \mathbf{1} \otimes (\mathbf{2}_L^0 \oplus \mathbf{2}_R^0)$$

$$1_3 \otimes P^q : \mathbf{3} \otimes (\mathbf{2}_L^0 \oplus \mathbf{2}_R^0) \to \mathbf{3} \otimes (\mathbf{2}_L^0 \oplus \mathbf{2}_R^0).$$

PROOF. The action of $\mathcal{A}_F$ on $\mathcal{H}_F = \mathcal{H}_f \oplus \mathcal{H}_{\bar{f}}$ is of the form

$$\begin{pmatrix} \pi(\lambda, q, m) & 0 \\ 0 & \pi'(\lambda, q, m) \end{pmatrix} \tag{1.673}$$

On the subspace $\mathcal{H}_f$ and in the decomposition on the basis $|\uparrow\rangle_R$, $|\downarrow\rangle_R$, $|\uparrow\rangle_L$, and $|\downarrow\rangle_L$ of $\mathbf{2}_L$ and $\mathbf{2}_R$, one has

$$\pi(\lambda, q, m) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \overline{\lambda} & 0 & 0 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\overline{\beta} & \overline{\alpha} \end{pmatrix} \otimes 1_{12}, \tag{1.674}$$

where $12 = \dim(\mathbf{1}^0 \oplus \mathbf{3}^0) \times N$ with $N = 3$. Since (1.673) is diagonal the condition $P \in \mathcal{A}'_F$ is expressed independently on the matrix elements P_{ij}.
Let us consider first the case of the element P_{11}. This must commute with operators of the form $\pi(\lambda, q, m) \otimes 1_{12}$ with π as in (1.674), and 1_{12} the unit matrix in a twelve-dimensional space. This means that the matrix of P_{11} is block diagonal with three blocks in $M_{12}(\mathbb{C})$, $M_{12}(\mathbb{C})$, and $1_2 \otimes M_{12}(\mathbb{C})$, corresponding to the subspaces where the action of (λ, q, m) is by λ, $\bar{\lambda}$ and q.
We consider next the case of P_{22}. The action of $(\lambda, q, m) \in \mathcal{A}_F$ in the subspace $\mathcal{H}_{\bar{f}}$ is given by multiplication by λ or by m. Thus, the only condition on P_{22} is that it is an operator of the form (1.672).
The off-diagonal terms P_{12} and P_{21} must intertwine the actions of $(\lambda, q, m) \in \mathcal{A}_F$ in $\mathcal{H}_f$ and $\mathcal{H}_{\bar{f}}$. However, the actions of q or m are disjoint in these two spaces, while only the action by λ occurs in both. The subspace of $\mathcal{H}_f$ on which (λ, q, m) acts by λ is $|\uparrow\rangle_R \otimes \mathbf{1}^0 \oplus |\uparrow\rangle_R \otimes \mathbf{3}^0$. The subspace of $\mathcal{H}_{\bar{f}}$ on which (λ, q, m) acts by λ is $\mathbf{1} \otimes \mathbf{2}_L^0 \oplus \mathbf{1} \otimes \mathbf{2}_R^0$. Thus, the conclusion follows from the intertwining condition. $\square$

We then obtain the classification of Dirac operators.

THEOREM 1.187. (1) *Let D be a Dirac operator for $(\mathcal{A}_F, \mathcal{H}_F, J_F, \gamma_F)$. Then there exist 3×3 matrices Υ_e, Υ_ν, Υ_d, Υ_u, and a symmetric 3×3 matrix Υ_R, such that D is of the form $D = D(\Upsilon)$ as in* (1.663).
(2) *All operators $D(\Upsilon)$ as in* (1.663) *with the matrix Υ_R of* (1.669) *symmetric are Dirac operators for the data $(\mathcal{A}_F, \mathcal{H}_F, J_F, \gamma_F)$.*
(3) *Two operators $D(\Upsilon)$ and $D(\Upsilon')$ are conjugate by a unitary operator commuting with $\mathcal{A}_F$, γ_F and J_F iff there exists unitary matrices V_j and W_j such that*

$$\begin{aligned} &\Upsilon'_e = V_1 \Upsilon_e V_3^*, \qquad \Upsilon'_\nu = V_2 \Upsilon_\nu V_3^*, \\ &\Upsilon'_d = W_1 \Upsilon_d W_3^*, \quad \Upsilon'_u = W_2 \Upsilon_u W_3^*, \\ &\Upsilon'_R = V_2 \Upsilon_R \bar{V}_2^*. \end{aligned} \tag{1.675}$$

PROOF. We first check (2), i.e. we show that operators of the form $D(\Upsilon)$ are Dirac operators provided Υ_R is symmetric. Let us first consider the off-diagonal part of $D(\Upsilon)$ in (1.663). This is of the form

$$\begin{pmatrix} 0 & \Upsilon_R^* \\ \Upsilon_R & 0 \end{pmatrix}.$$

Anticommutation with γ_F holds, since the operator γ_F restricted to the space $E_R \oplus J_F E_R$, with E_R as in (1.670), is of the form

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Moreover, the off-diagonal part of $D(\Upsilon)$ commutes with J_F iff

$$\overline{(\Upsilon_R \xi)} = \Upsilon_R^* \bar{\xi}, \qquad \forall \xi, \tag{1.676}$$

i.e. iff Υ_R is a symmetric matrix. The order one condition is automatically satisfied, since the commutator with elements of $\mathcal{A}_F$ vanishes identically.

We can now consider the diagonal part

$$\begin{pmatrix} S & 0 \\ 0 & \bar{S} \end{pmatrix}$$

of $D(\Upsilon)$. It commutes with J_F and anticommutes with γ_F by construction. It is then sufficient to check the commutation with $\mathbb{C}_F \subset \mathcal{A}_F$ and the order one condition on the subspace $\mathcal{H}_f$. Since S exactly commutes with the action of $\mathcal{A}_F^0$, the order one condition follows. In fact, for any $b \in \mathcal{A}_F$, the action of b^0 commutes with any operator of the form (1.672). This yields the order one condition, since $P = [S, \pi(a)]$ is of this form. The action of $\mathcal{A}_F$ on the subspace $\mathcal{H}_F$ is given by (1.674) and one checks that $\pi(\lambda, \lambda, 0)$ commutes with S, since the matrix of S has no non-zero element intertwining the $|\uparrow\rangle$ and the $|\downarrow\rangle$ subspaces.

We now show (1), i.e. that all Dirac operators are indeed of the form $D(\Upsilon)$. Let D be a Dirac operator. Since D is self-adjoint and commutes with J_F it is of the form

$$D = \begin{pmatrix} S & T^* \\ T & \bar{S} \end{pmatrix}$$

with $T = T^t$ symmetric.

Consider the element $v = (-1, 1, 1) \in \mathcal{A}_F$. One has

$$\gamma_F \, \xi = \, v \, \xi \, , \quad \forall \xi \in \mathcal{H}_f. \tag{1.677}$$

Notice that this equality fails on $\mathcal{H}_{\bar{f}}$.

The anticommutation of D with γ_F implies that $D = -\frac{1}{2} \gamma_F \, [D, \gamma_F]$. Notice that γ_F is given by a diagonal matrix of the form

$$\gamma_F = \begin{pmatrix} g & 0 \\ 0 & -\bar{g} \end{pmatrix}.$$

Thus, using (1.677), we get

$$S = -\frac{1}{2} \, g \, [S, g] = -\frac{1}{2} \, v \, [S, v].$$

The action of v in $\mathcal{H}_F$ is given by a diagonal matrix (1.673), hence $v \, [S, v]$ coincides with the A_{11} block of the matrix of $A = v[D, v]$.

Thus, the order one condition implies that S commutes with all operators b^0, hence that it is of the form (1.664).

The anticommutation with γ_F and the commutation with $\mathbb{C}_F$ then imply that the self-adjoint matrix S can be written in the form of (1.666) and (1.668).

It remains to determine the form of the matrix T. The conditions on the off-diagonal elements of a matrix

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}$$

ensuring that P belongs to the commutant of $\mathcal{A}_F^0 = J_F \mathcal{A}_F J_F$, are the following.

- P_{12} has support in $\mathbf{1} \otimes |\uparrow\rangle_R^0 \oplus \mathbf{3} \otimes |\uparrow\rangle_R^0$ and range in $\mathbf{2}_L \otimes \mathbf{1}^0 \oplus \mathbf{2}_R \otimes \mathbf{1}^0$.
- P_{21} has support in $\mathbf{2}_L \otimes \mathbf{1}^0 \oplus \mathbf{2}_R \otimes \mathbf{1}^0$ and range in $\mathbf{1} \otimes |\uparrow\rangle_R^0 \oplus \mathbf{3} \otimes |\uparrow\rangle_R^0$.

This follows from Lemma 1.186, using J_F.
Consider then the element $e = (0,1,0)$ in $\mathcal{A}_F$. One has $\pi'(e) = 0$ and $\pi(e)$ is the projection onto the eigenspace $\gamma_F = 1$ in $\mathcal{H}_F$. Thus, since $[D, e]$ belongs to the commutant of $\mathcal{A}_F^0 = J_F \mathcal{A}_F J_F$ by the order one condition, one gets that $\pi'(e)T - T\pi(e) = -T\pi(e)$ has support in $\mathbf{2}_L \otimes \mathbf{1}^0 \oplus \mathbf{2}_R \otimes \mathbf{1}^0$ and range in $\mathbf{1} \otimes |\uparrow\rangle_R^0 \oplus \mathbf{3} \otimes |\uparrow\rangle_R^0$. In particular $\gamma_F = 1$ on the range.
Thus, the anticommutation with γ_F shows that the support of $T\pi(e)$ is in the eigenspace $\gamma_F = -1$, so that $T\pi(e) = 0$.
Consider then the element $e_3 = (0,0,1) \in \mathcal{A}_F$. Let us show that $Te_3^0 = 0$. By Definitions 1.120 and 1.124, a Dirac operator is a self-adjoint operator D in $\mathcal{H}_F$ commuting with J_F, $\mathbb{C}_F$, anticommuting with γ_F and fulfilling the order one condition $[[D,a],b^0] = 0$ for any $a, b \in \mathcal{A}_F$. Thus, T commutes with the actions of $v(\lambda) = (\lambda, \lambda, 0) \in \mathcal{A}_F$ and of $J_F v(\lambda) J_F^{-1} = v(\lambda)^0$. Thus, it commutes with e_3^0. The action of e_3^0 on $\mathcal{H}_f$ is the projection onto the subspace $\bullet \otimes \mathbf{3}^0$. The action of e_3^0 on $\mathcal{H}_{\bar{f}}$ is zero. Thus, $[T, e_3^0] = T\, e_3^0$ is the restriction of T to the subspace $\bullet \otimes \mathbf{3}^0$. Since $[T, e_3^0] = 0$ we get $T\, e_3^0 = 0$. We have shown that the support of T is contained in $\mathbf{2}_R \otimes \mathbf{1}^0$. Since T is symmetric, i.e. $T = \bar{T}^*$, the range of T is contained in $\mathbf{1} \otimes \mathbf{2}_R^0$. The left and right actions of (λ, q, m) on these two subspaces coincide with the left and right actions of $v(\lambda)$. Thus, we get that T commutes with $\mathcal{A}_F$ and $\mathcal{A}_F^0$. Thus, by Lemma 1.186, it has support in $|\uparrow\rangle_R \otimes \mathbf{1}^0$ and range in $\mathbf{1} \otimes |\uparrow\rangle_R^0$.
This means that T is given by a symmetric 3×3 matrix Υ_R, and the operator D is of the form $D = D(\Upsilon)$.

3) By Lemma 1.186, the commutant of the algebra generated by $\mathcal{A}_F$ and $\mathcal{A}_F^0$ is the algebra of matrices

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}$$

such that

- P_{12} has support in $\mathbf{1} \otimes |\uparrow\rangle_R^0$ and range in $|\uparrow\rangle_R \otimes \mathbf{1}^0$.
- P_{21} has support in $|\uparrow\rangle_R \otimes \mathbf{1}^0$ and range in $\mathbf{1} \otimes |\uparrow\rangle_R^0$.
- P_{jj} is of the form

  $$P_{jj} = P_{jj}^{\ell} \oplus (P_{jj}^{q} \otimes 1_3)$$

  where

  $$P_{jj}^{a} = \begin{pmatrix} P_j^a(1) & 0 & 0 \\ 0 & P_j^a(2) & 0 \\ 0 & 0 & 1_2 \otimes P_j^a(3) \end{pmatrix}.$$

  for $j = 1, 2$ and for $a = \ell$ or $a = q$.

A unitary operator U acting on $\mathcal{H}_F$ commuting with $\mathcal{A}_F$ and J is in the commutant of the algebra generated by $\mathcal{A}_F$ and $\mathcal{A}_F^0$. If it commutes with γ_F, then the off-diagonal elements U_{ij} vanish, since $\gamma_F = -1$ on $|\uparrow\rangle_R \otimes \mathbf{1}^0$ and $\gamma_F = 1$ on $\mathbf{1} \otimes |\uparrow\rangle_R^0$.

Thus, U is determined by the six 3×3 matrices $U_1^a(k)$. In fact, since it commutes with J_F, one has $U_2^a(k) = \bar{U}_1^a(k)$. One checks that conjugating by U gives part (3) of the theorem. □

Notice that it is crucial in order to have the term Υ_R in the Dirac operator D_F that γ_F gives opposite signs to the corresponding chiralities for particles and antiparticles, i.e. that it anticommutes rather than commutes with the real structure J_F. This anticommutation is, as we observed before, the sign that the KO-dimension is different from the metric dimension and is 6 instead of 0 modulo 8. Physically, the term Υ_R is the one that gives the Majorana mass term for neutrinos and accounts for the see-saw mechanism for neutrino masses, as we discuss in §17.5.

A first consequence of physical significance that one can derive from the classification result of Theorem 1.187 is that, in our model, color is unbroken, which is in agreement with the physics of the Standard Model. This is a direct consequence of the fact that all Dirac operators D_F for the finite geometry $(\mathcal{A}_F, \mathcal{H}_F, J_F, \gamma_F)$ are of the form (1.663), with the S term of the form (1.664), with the part of S acting on the quark sector of the form $S_q \otimes 1_3$, with the identity matrix on the color indices.

In terms of the general form of Dirac operators described in Theorem 1.187 above, we can also characterize the subalgebra $\mathcal{A}_F$ of the left-right symmetric algebra $\mathcal{A}_{LR}$ that was our initial input through the following result.

COROLLARY 1.188. *Let $D_F^{\rm off}$ denote the off-diagonal part of a Dirac operator $D_F = D(\Upsilon)$ as in* (1.663)*, with the term $\Upsilon_R \neq 0$. Then the algebra $\mathcal{A}_F \subset \mathcal{A}_{LR}$ is characterized by the property*

$$\mathcal{A}_F = \{a \in \mathcal{A}_{LR} \mid [D_F^{\rm off}, a] = 0\}. \tag{1.678}$$

PROOF. Let $a = (\lambda, q_L, q_R, m) \in \mathcal{A}_{LR}$. Let D_F be an operator of the form $D_F = D(\Upsilon)$ as in (1.663). By construction $D_F^{\rm off}$ has support in $E_R \oplus J_F E_R$. The action of a on $J_F E_R$ is given by multiplication by λ. The commutation $[D_F^{\rm off}, a] = 0$ shows that this action must agree with the action of a on E_R. The latter is given by multiplication by a quaternion $q_R = \alpha + \beta j$

$$a\,\nu_R = \alpha \nu_R - \bar{\beta} e_R,$$

where we identify the basis with the fermions as in §13.2. Thus, $[D_F^{\rm off}, a] = 0$ implies that $\beta = 0$ and $\alpha = \lambda$, hence that $q_R = \lambda$ and $a \in \mathcal{A}_F$. Since we already saw that $[D_F^{\rm off}, a] = 0$ for all $a \in \mathcal{A}_F$ we have obtained the required equality. □

Physically, the result of Corollary 1.188 means that the presence of a non-zero Majorana mass term Υ_R for the neutrinos does not generate new boson fields. This will be clearer in §15.2 where we show how the Higgs bosons arise from the inner fluctuations of the metric.

We now describe geometrically the moduli space of Dirac operators D_F, using the classification of Theorem 1.187.

13.5. Moduli space of Dirac operators and Yukawa parameters.

We introduce the following notation.

DEFINITION 1.189. *Let $\mathcal{C}_q$ denote the moduli space of pairs of invertible 3×3 matrices (Υ_d, Υ_u) modulo the equivalence relation*

$$\Upsilon'_d = W_1\,\Upsilon_d\,W_3^*, \qquad \Upsilon'_u = W_2\,\Upsilon_u\,W_3^*, \tag{1.679}$$

where the W_j are unitary matrices. Let $\mathcal{C}_\ell$ denote the moduli space of triplets of 3×3 matrices $(\Upsilon_e, \Upsilon_\nu, \Upsilon_R)$, with Υ_R symmetric, modulo the equivalence relation

$$\Upsilon'_e = V_1\,\Upsilon_e V_3^* \qquad \Upsilon'_\nu = V_2\,\Upsilon_\nu\,V_3^* \tag{1.680}$$

and

$$\Upsilon'_R = V_2\,\Upsilon_R\,\bar{V}_2^*. \tag{1.681}$$

We also need the following preliminary results that will be used in Proposition 1.192 to give an explicit description of the moduli space of Dirac operators D_F.

LEMMA 1.190. *Consider a fixed basis for the space $\mathcal{H}_F$ as in §13.2. Each equivalence class of pairs of 3×3 matrices under the relation (1.679) contains a pair (Υ_d, Υ_u) where Υ_u is diagonal in the given basis and has positive entries, while Υ_d is positive and of the form $C\delta_\downarrow C^*$, with $\delta_\downarrow$ diagonal and $C \in \mathrm{SU}(3)$.*

PROOF. We can use the freedom to choose W_2 and W_3 to make Υ_u positive and diagonal and the freedom to choose W_1 to make Υ_d positive.
The eigenvalues are the characteristic values (i.e. the eigenvalues of the absolute value in the polar decomposition) of Υ_u and Υ_d and are invariants of the pair.
Thus, we can find diagonal matrices $\delta_\uparrow$ and $\delta_\downarrow$ and a unitary matrix C such that

$$\Upsilon_u = \delta_\uparrow \quad \text{and} \quad \Upsilon_d = C\,\delta_\downarrow\,C^*. \tag{1.682}$$

Since multiplying C by a scalar does not affect the result, we can assume that $\det(C) = 1$. □

Thus, $C \in \mathrm{SU}(3)$ depends a priori upon 8 real parameters. However, only the double coset of C modulo the diagonal subgroup $\mathcal{N} \subset \mathrm{SU}(3)$ matters, by the following result.

LEMMA 1.191. *Suppose given a fixed pair of diagonal matrices $\delta_\uparrow$ and $\delta_\downarrow$ with positive and distinct eigenvalues. Let $\mathcal{N} \subset \mathrm{SU}(3)$ denote the subgroup of diagonal matrices. Suppose given two matrices C and C' in $\mathrm{SU}(3)$ and consider the corresponding pairs as in Lemma 1.191 of the form*

$$(\delta_\uparrow, C\,\delta_\downarrow\,C^*), \qquad (\delta_\uparrow, C'\,\delta_\downarrow\,C'^*). \tag{1.683}$$

The pairs in (1.683) are equivalent iff there exists diagonal unitary matrices $A, B \in \mathcal{N}$ such that

$$AC = C'B. \tag{1.684}$$

PROOF. For $AC = C'B$ one has

$$(A\delta_\uparrow A^*, AC\delta_\downarrow C^*A^*) = (\delta_\uparrow, C'B\delta_\downarrow B^*C'^*) = (\delta_\uparrow, C'\delta_\downarrow C'^*), \tag{1.685}$$

since A and B are unitary and diagonal, hence the two pairs are equivalent. Conversely, with W_j as in (1.679) one gets $W_1 = W_3$ from the uniqueness of the polar decomposition

$$\delta_\downarrow = (W_1W_3^*)\,(W_3\delta_\downarrow W_3^*).$$

Similarly, one obtains $W_2 = W_3$. Thus, $W_3 = W$ is diagonal and we get

$$W\,C\,\delta_\downarrow\,C^*\,W^* = C'\,\delta_\downarrow\,C'^*,$$

so that $WC = C'B$ for some diagonal matrix B. Since W and B have the same determinant one can assume that they both belong to $\mathcal{N}$. $\square$

We then have the following explicit description of the moduli space of Dirac operators D_F.

PROPOSITION 1.192. *In the case of $N = 3$ generations, the moduli space of Dirac operators D_F for the geometry $(\mathcal{A}_F, \mathcal{H}_F, J_F, \gamma_F)$ is the real 31-dimensional space*

$$\mathcal{C}_\ell \times \mathcal{C}_q. \tag{1.686}$$

The space $\mathcal{C}_q$ is of real dimension 10 and given by the double coset space

$$\mathcal{C}_q \cong (\mathrm{U}(3) \times \mathrm{U}(3)) \backslash (\mathrm{GL}_3(\mathbb{C}) \times \mathrm{GL}_3(\mathbb{C}))/\mathrm{U}(3), \tag{1.687}$$

where $\mathrm{U}(3)$ acts diagonally on the right. The space $\mathcal{C}_\ell$ is of real dimension 21 and fibers over $\mathcal{C}_q$, with generic fiber the quotient of symmetric complex 3×3 matrices by $\mathrm{U}(1)$. In fact, $\mathcal{C}_\ell$ is given by the quotient

$$\mathcal{C}_\ell \cong (\mathrm{U}(3) \times \mathrm{U}(3)) \backslash (\mathrm{GL}_3(\mathbb{C}) \times \mathrm{GL}_3(\mathbb{C}) \times \mathcal{S})/\mathrm{U}(3), \tag{1.688}$$

with $\mathcal{S}$ the space of symmetric complex 3×3 matrices. The action of $\mathrm{U}(3) \times \mathrm{U}(3)$ on the left in (1.688) is given by left multiplication on $\mathrm{GL}_3(\mathbb{C}) \times \mathrm{GL}_3(\mathbb{C})$ and by (1.681) on $\mathcal{S}$. The action of $\mathrm{U}(3)$ on the right in (1.688) is trivial on $\mathcal{S}$ and by diagonal right multiplication on $\mathrm{GL}_3(\mathbb{C}) \times \mathrm{GL}_3(\mathbb{C})$.

PROOF. By Lemma 1.190 we see that the real dimension of the moduli space $\mathcal{C}_q$ is $3+3+4 = 10$, where the $3+3$ comes from the eigenvalues and the $4 = 8-4$ from the double coset space of C's in $\mathrm{SU}(3)$ modulo $\mathcal{N}$ as in Lemma 1.191. One way to parameterize the representatives of the double cosets of the matrix C is by means of three angles θ_i and a phase δ, in the form (1.419) where $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, and $e_\delta = \exp(i\delta)$. One has by construction the factorization

$$C = R_{23}(\theta_2)\, d(\delta)\, R_{12}(\theta_1)\, R_{23}(-\theta_3) \tag{1.689}$$

where $R_{ij}(\theta)$ is the rotation of angle θ in the ij-plane and $d(\delta)$ the diagonal matrix

$$d(\delta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -e^{i\delta} \end{pmatrix}.$$

The identification of $\mathcal{C}_q$ with the double coset (1.687) follows from the explicit form of the equivalence relation (1.679).

Consider then the moduli space $\mathcal{C}_\ell$. By construction one has a natural surjective map

$$\pi : \mathcal{C}_\ell \to \mathcal{C}_q \tag{1.690}$$

given by forgetting the Υ_R term. The generic fiber of π is the space of symmetric complex 3×3 matrices modulo the action of a complex scalar λ of absolute value one given by

$$\Upsilon_R \mapsto \lambda^2\, \Upsilon_R. \tag{1.691}$$

The (real) dimension of the fiber is $12 - 1 = 11$. The total real dimension of the moduli space $\mathcal{C}_\ell$ is then 21. The identification with the space (1.688) again follows from the explicit form of the equivalence relation (1.679) together with (1.680).

It then follows from Theorem 1.187 and the results above on $\mathcal{C}_\ell$ and $\mathcal{C}_q$ that the total moduli space of Dirac operators D_F is the product $\mathcal{C}_\ell \times \mathcal{C}_q$, which is of real dimension 31. ☐

The 31 real parameters of (1.686) account for all the Yukawa parameters in the Standard Model with neutrino mixing and Majorana mass terms. In fact, the parameters in $\mathcal{C}_3$ correspond to the masses of the quarks and the quark mixing angles of the CKM matrix, while the additional parameters of $\mathcal{C}_1$ give the lepton masses, the angles of the PMNS mixing matrix and the Majorana mass terms for neutrinos.
This gives a clear geometric meaning to the Yukawa parameters as the choice of a point in $\mathcal{C}_\ell \times \mathcal{C}_q$, i.e. of a Dirac operator D_F for the finite geometry $(\mathcal{A}_F, \mathcal{H}_F, J_F, \gamma_F)$. It is interesting to notice that the space $\mathcal{C}_\ell \times \mathcal{C}_q$ admits interesting compactifications, possibly including something akin to the noncommutative compactifications of the modular curves considered in [**77**] and [**219**]. We return to this observation at the very end of the book when spectral correspondences between different geometries are introduced in the attempt to illustrate a parallel between the problem of quantum gravity and the noncommutative compactifications that we present in Chapter 3 in the number theoretic context related to the Riemann zeta function.

13.6. The intersection pairing of the finite geometry.

The fact that, from the point of view of KO-homology, the dimension of the finite geometry $(\mathcal{A}_F, \mathcal{H}_F, D_F)$ is equal to 6 modulo 8 implies that the intersection pairing is *skew-symmetric*. It is given explicitly as follows.

PROPOSITION 1.193. *The expression*

$$\langle e, f\rangle = \mathrm{Tr}(\gamma_F e\, J_F f J_F^{-1}) \tag{1.692}$$

defines an antisymmetric bilinear pairing on $K_0(\mathcal{A}_F) \times K_0(\mathcal{A}_F)$. *The group* $K_0(\mathcal{A}_F)$ *is the free abelian group generated by the classes of* $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ *and* $f_3 = (0,0,f)$, *where* $f \in M_3(\mathbb{C})$ *is a minimal idempotent.*

PROOF. The pairing (1.692) is obtained from the composition of the natural map

$$K_0(\mathcal{A}_F) \times K_0(\mathcal{A}_F) \to K_0(\mathcal{A}_F \otimes \mathcal{A}_F^0)$$

with the graded trace $\mathrm{Tr}(\gamma_F\,\cdot)$. Since J_F anticommutes with γ_F, one checks that

$$\begin{aligned}\langle f, e\rangle = &\ \ \mathrm{Tr}(\gamma_F f J_F e J_F^{-1}) = -\mathrm{Tr}(\gamma_F J_F^{-1} f J_F e)\\ = &\ -\mathrm{Tr}(\gamma_F e J_F f J_F^{-1}) = -\langle e, f\rangle,\end{aligned}$$

i.e. that the pairing is antisymmetric.
By construction, $\mathcal{A}_F$ is the direct sum of the field $\mathbb{C}$, the division algebra $\mathbb{H}$ and the algebra $M_3(\mathbb{C})$. The latter is Morita equivalent to $\mathbb{C}$. The projections $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ and $f_3 = (0,0,f)$ are the three minimal idempotents in $\mathcal{A}_F$. ☐

DEFINITION 1.194. *Let* $\mathcal{H}_{F,\ell}$ *and* $\mathcal{H}_{F,q}$ *be given by*

$$\mathcal{H}_{F,\ell} := \mathbf{2}_L \otimes \mathbf{1}^0 \oplus \mathbf{2}_R \otimes \mathbf{1}^0 \oplus \mathbf{1} \otimes \mathbf{2}_L^0 \oplus \mathbf{1} \otimes \mathbf{2}_R^0 \tag{1.693}$$

and

$$\mathcal{H}_{F,q} := \mathbf{2}_L \otimes \mathbf{3}^0 \oplus \mathbf{2}_R \otimes \mathbf{3}^0 \oplus \mathbf{3} \otimes \mathbf{2}_L^0 \oplus \mathbf{3} \otimes \mathbf{2}_R^0. \tag{1.694}$$

By construction, the KO-homology class $(\mathcal{H}_F, \gamma_F, J_F)$ given by the representation of $\mathcal{A}_F$ on $\mathcal{H}_F$ with the $\mathbb{Z}/2$-grading γ_F and the real structure J_F splits as a direct sum of two pieces, corresponding respectively to leptons and quarks.

PROPOSITION 1.195. (1) *The representation of the algebra generated by* $(\mathcal{A}_F, D_F, J_F, \gamma_F)$ *on* $\mathcal{H}_F$ *splits as a direct sum of two subrepresentations*

$$\mathcal{H}_F = \mathcal{H}_{F,\ell} \oplus \mathcal{H}_{F,q}, \tag{1.695}$$

with $\mathcal{H}_{F,\ell}$ *and* $\mathcal{H}_{F,q}$ *as in Definition 1.194.*

(2) *In the generic case where the matrices* Υ_ν, Υ_e, Υ_u, Υ_d *and* Υ_R *in* D_F *have distinct eigenvalues, each of these subrepresentations is irreducible.*

(3) *In the basis* (e_1, e_2, f_3) *the pairing* (1.692) *is given by*

$$\langle \cdot, \cdot \rangle_\ell = \begin{pmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \langle \cdot, \cdot \rangle_q = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ -2 & 2 & 0 \end{pmatrix}, \tag{1.696}$$

up to an overall multiplicity given by a factor of $N = 3$ *that corresponds to the number of generations.*

PROOF. (1) By construction, the action of $\mathcal{A}_F$ in $\mathcal{H}_F$ is block diagonal in the decomposition $\mathcal{H}_F = \mathcal{H}_{F,\ell} \oplus \mathcal{H}_{F,q}$. Both the actions of J_F and of γ_F are also block diagonal. Theorem 1.187 shows that D_F is also block diagonal, since it is of the form $D = D(\Upsilon)$.

(2) It is enough to show that a unitary operator that commutes with $\mathcal{A}_F$, γ_F, J_F and D_F is a scalar. Let us start with $\mathcal{H}_{F,q}$. By (3) of Theorem 1.187, such a unitary operator is given by three unitary matrices $W_j \in M_3(\mathbb{C})$ such that

$$\Upsilon_d = W_1 \, \Upsilon_d \, W_3^*, \qquad \Upsilon_u = W_2 \, \Upsilon_u \, W_3^*.$$

We can assume that both Υ_u and Υ_d are positive and that Υ_u is diagonal, by Lemma 1.190.

The uniqueness of the polar decomposition then shows that

$$\Upsilon_d = (W_1 W_3^*) \, (W_3 \Upsilon_d W_3^*) \quad \Rightarrow \quad W_1 W_3^* = 1 \quad \text{and} \quad W_3 \Upsilon_d W_3^* = \Upsilon_d.$$

Thus, we get $W_1 = W_2 = W_3$. Since generically all the eigenvalues of Υ_u or Υ_d are distinct, we get that the matrices W_j are diagonal in the basis of eigenvectors of the matrices Υ_u and Υ_d. However, generically these bases are distinct, hence we conclude that $W_j = 1$ for all j. The same result holds "a fortiori" for $\mathcal{H}_{F,\ell}$ where the conditions imposed by (3) of Theorem 1.187 are in fact stronger.

(3) One computes the pairing directly using the definition of γ_F. On $\mathcal{H}_{F,\ell}$ the subalgebra $M_3(\mathbb{C})$ acts by zero which explains why the last line and columns of the pairing matrix vanish. By antisymmetry one just needs to evaluate

$$\langle e_1, e_2 \rangle = -\langle e_2, e_1 \rangle = -\mathrm{Tr}(\gamma_F \, e_2 \, J_F e_1 J_F^{-1}) = -\mathrm{Tr}(\gamma_F \, e_2) = 2 \times 3,$$

where 3 is the number of generations. In fact, by the explicit form (1.652) of the action of the algebra on leptons and quarks one sees that e_2 acts by zero except on the piece $\mathbf{2}_L \otimes \mathbf{1}^0$ of $\mathcal{H}_{F,\ell}$ where it acts as the identity. On $\mathcal{H}_{F,q}$ the same pair gives $\langle e_1, e_2 \rangle = 0$, since now the right action of e_1 is zero on $\mathcal{H}_f$. In the same way one gets $\langle f_3, e_2 \rangle = 2 \times 3$. Finally one has, by the same argument,

$$\langle e_1, f_3 \rangle = \mathrm{Tr}(\gamma_F e_1 J_F f_3 J_F^{-1}) = 2 \times 3.$$

□

An antisymmetric 3×3 matrix is automatically degenerate since its determinant vanishes. The following result shows that one can still obtain a nondegenerate bilinear map with values not in $\mathbb{R}$ but in $\mathbb{R}^2$. This corresponds to the fact that the KO-homology is not singly generated as a module over K_0, but it is generated by two elements.

COROLLARY 1.196. *The pairing* $K_0(\mathcal{A}_F) \oplus K_0(\mathcal{A}_F) \to \mathbb{R} \oplus \mathbb{R}$ *given by*

$$\langle \cdot, \cdot \rangle_F := \langle \cdot, \cdot \rangle_\ell \oplus \langle \cdot, \cdot \rangle_q \tag{1.697}$$

is non-degenerate.

PROOF. We need to check that, for any e in $K_0(\mathcal{A}_F)$ there exists an $f \in K_0(\mathcal{A}_F)$ such that $\langle e, f \rangle_F \neq (0,0)$. This can be seen by the explicit form of $\langle \cdot, \cdot \rangle_\ell$ and $\langle \cdot, \cdot \rangle_q$ in (1.696). □

14. The product geometry

In the previous section we described the finite geometry

$$F = (\mathcal{A}_F, \mathcal{H}_F, D_F, J_F, \gamma_F).$$

We recovered the fermions of the Standard Model with the correct hypercharge assignment and gave a geometric interpretation for the Yukawa parameters (including the Majorana mass terms for neutrinos) using the classification of Dirac operators D_F. Now we consider the product of the finite noncommutative geometry F with the spectral triple associated to the commutative geometry of a compact 4-dimensional Riemannian spin manifold of space-time (in Euclidean signature).

Recall that, for M a compact spin 4-manifold, the associated spectral triple is given by $(C^\infty(M), L^2(M,S), \partial\!\!\!/_M)$, where $L^2(M,S)$ is the space of square integrable spinors, with smooth functions in $C^\infty(M)$ acting as multiplication operators, and with $\partial\!\!\!/_M$ the Dirac operator. The grading γ_M is given by the chirality operator, which is usually denoted by γ_5 in the 4-dimensional case. The operator J_M that gives the real structure is the *charge conjugation operator* (cf. (1.407)). We refer to [**132**] for a detailed treatment of spectral triples associated to Riemannian manifolds.

Suppose given two real spectral triples

$$(\mathcal{A}_1, \mathcal{H}_1, D_1, J_1, \gamma_1) \quad \text{and} \quad (\mathcal{A}_2, \mathcal{H}_2, D_2, J_2, \gamma_2),$$

where the first is of KO-dimension 4 modulo 8 and the second is of KO-dimension 6 modulo 8. Then the product geometry is a real spectral triple

$$(\mathcal{A}, \mathcal{H}, D, J, \gamma) = (\mathcal{A}_1, \mathcal{H}_1, D_1, J_1, \gamma_1) \otimes (\mathcal{A}_2, \mathcal{H}_2, D_2, J_2, \gamma_2) \tag{1.698}$$

with

$$\begin{gathered} \mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2, \quad \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2, \quad D = D_1 \otimes 1 + \gamma_1 \otimes D_2, \\ \gamma = \gamma_1 \otimes \gamma_2, \quad J = J_1 \otimes J_2. \end{gathered} \tag{1.699}$$

The resulting geometry $(\mathcal{A}, \mathcal{H}, D, J, \gamma)$ is of KO-dimension $10 = 2$ modulo 8. Notice that it matters here that J_1 commutes with γ_1, in order to check that J commutes with D. One checks that the order one condition is fulfilled by D if it is fulfilled by both D_1 and D_2. We refer to §9.5 of [**132**] for a thorough discussion of the notions of KO-dimension and products of spectral triples.

DEFINITION 1.197. *The product $M \times F$ of a compact spin 4-manifold with the finite noncommutative geometry F is the real spectral triple*

$$(\mathcal{A}, \mathcal{H}, D, J, \gamma) = (C^\infty(M), L^2(M, S), \partial\!\!\!/_M, J_M, \gamma_5) \otimes (\mathcal{A}_F, \mathcal{H}_F, D_F, J_F, \gamma_F) \tag{1.700}$$

defined as in (1.699).

14.1. The real part of the product geometry.

As in Proposition 1.125 of §10.2, we can associate to a real spectral triple the commutative geometry given by its *real part.* In the case of the product geometry $M \times F$ of Definition 1.197 we obtain the following results.

LEMMA 1.198. *Let F be the finite noncommutative geometry*

$$F = (\mathcal{A}_F, \mathcal{H}_F, D_F, J_F, \gamma_F)$$

and let $M \times F$ be the product with a spin 4-manifold as in Definition 1.197. Then the following holds.

- *The real part of F is given by*

$$(\mathcal{A}_F)_{J_F} = \mathbb{R} = \{(\lambda, \lambda, \lambda) \mid \lambda \in \mathbb{R}\} \subset \mathcal{A}_F. \tag{1.701}$$

- *The real part $\mathcal{A}_J = C^\infty(M, \mathcal{A}_F)_{J_M \otimes J_F}$ of the product geometry $M \times F$ is*

$$\mathcal{A}_J = C^\infty(M, \mathbb{R}). \tag{1.702}$$

PROOF. Let $a = (\lambda, q, m) \in \mathcal{A}_F$. Then, if a commutes with J_F, its action in $\mathcal{H}_f \subset \mathcal{H}_F$ coincides with the right action of a^*. When considering the action on $\mathcal{H}_{F,\ell}$, this implies that $\lambda = \bar{\lambda}$ and that the action of the quaternion $q \in \mathbb{H}$ coincides with that of λ. Thus, one has $q = \lambda \in \mathbb{R}$. When looking at the action on $\mathcal{H}_{F,q}$, one obtains similarly that $m = \lambda$. The same proof applies to $C^\infty(M, \mathcal{A}_F)$. $\square$

The above result shows that the real part of the product geometry only sees the ordinary 4-dimensional commutative space-time M as a real manifold.

15. Bosons as inner fluctuations

In this section we show how to obtain the bosons of the Standard Model, with the correct quantum numbers, as inner fluctuations of the metric for the product geometry $M \times F$.

We begin by showing that, as remarked in §9.9, the symmetries of the geometry $M \times F$ recover the group of local gauge transformations of the Standard Model prior to symmetry breaking, that is, the semidirect product $\mathcal{G}_{SM} \rtimes \mathrm{Diff}(M)$ of (1.465), with $\mathcal{G}_{SM} = \mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)$.

We then compute the inner fluctuations of the metric for $M \times F$. We separate the computation into two parts. In fact, consider elements of the form $A = \sum_i a_i [D, a_i']$, with $a_i, a_i' \in \mathcal{A}$. Since the Dirac operator for the product geometry is of the form $D = \partial\!\!\!/_M \otimes 1 + \gamma_5 \otimes D_F$, the elements A as above correspondingly decompose as a sum of two terms $A = A^{(0,1)} + A^{(1,0)}$ (notation not to be confused with a Hodge decomposition). We first consider the "discrete part" $A^{(0,1)}$ coming from commutators with $\gamma_5 \otimes D_F$ and then the "continuous part" $A^{(1,0)}$ coming from $\partial\!\!\!/_M \otimes 1$. We show that the first part generates the Higgs field, while the second part produces all the other bosons for the Standard Model.

15.1. The local gauge transformations.

The following result shows how to recover the local gauge symmetries of the Standard Model minimally coupled to gravity from the group

$$C^\infty(M, \mathrm{SU}(\mathcal{A}_F))$$

of symmetries of the product geometry $M \times F$.

PROPOSITION 1.199. *Let $(\mathcal{A}, \mathcal{H}, D)$ be the real spectral triple associated to $M \times F$.*

- *Let U be a unitary operator on $\mathcal{H}$, which commutes with γ and J, and such that $U\,\mathcal{A}\,U^* = \mathcal{A}$. Then there exists a unique diffeomorphism $\varphi \in \mathrm{Diff}(M)$ such that*

$$U\,f\,U^* = f \circ \varphi \qquad \forall f \in \mathcal{A}_J. \tag{1.703}$$

- *Let U be a unitary operator as above, such that the associated diffeomorphism is trivial, $\varphi = \mathrm{id}$. Then, possibly after passing to a finite abelian cover of M, there exists a unitary $u \in C^\infty(M, \mathrm{SU}(\mathcal{A}_F))$ such that $U\,\mathrm{Ad}(u)^* \in \mathcal{C}$, where $\mathcal{C}$ is the commutant of the algebra of operators on $\mathcal{H}$ generated by $\mathcal{A}$ and $J\mathcal{A}J^{-1}$.*

PROOF. We refer to [**204**] for finer points concerning the lifting of diffeomorphisms preserving the given spin structure.
The first statement follows from the functoriality of the construction of the subalgebra $\mathcal{A}_J$ and the classical result that automorphisms of the algebra $C^\infty(M, \mathbb{R})$ are given by composition with a diffeomorphism of M.
Let us prove the second statement. One has $\mathcal{H} = L^2(M, S) \otimes \mathcal{H}_F = L^2(M, S \otimes \mathcal{H}_F)$. Since $\varphi = \mathrm{id}$, we know by (1.703) that U commutes with the algebra $\mathcal{A}_J = C^\infty(M, \mathbb{R})$. This shows that U is given by an endomorphism $x \mapsto U(x)$ of the vector bundle $S \otimes \mathcal{H}_F$ on M. Since U commutes with J, the unitary $U(x)$ commutes with $J_x \otimes J_F$, where J_x is the restriction of J_M to the fiber.
The equality $U\,\mathcal{A}\,U^* = \mathcal{A}$ shows that, for all $x \in M$, one has

$$U(x)\,(\mathrm{id} \otimes \mathcal{A}_F)\,U^*(x) = \mathrm{id} \otimes \mathcal{A}_F\,. \tag{1.704}$$

Here we identify $\mathcal{A}_F$ with a subalgebra of operators on $S \otimes \mathcal{H}_F$, through the algebra homomorphism $T \mapsto \mathrm{id} \otimes T$.
Let α be an arbitrary automorphism of $\mathcal{A}_F$. The center of $\mathcal{A}_F$ contains three minimal idempotents and the corresponding reduced algebras $\mathbb{C}$, $\mathbb{H}$, $M_3(\mathbb{C})$ are pairwise non-isomorphic. Thus α preserves these three idempotents and is determined by its restriction to the corresponding reduced algebras $\mathbb{C}$, $\mathbb{H}$, $M_3(\mathbb{C})$. In particular, such an automorphism will act on the subalgebra $\mathbb{C}$ either as the identity or as complex conjugation.
Now consider the automorphism α_x of $\mathcal{A}_F$ determined by (1.704). It is unitarily implemented by (1.704). The action of $\mathbb{C} \subset \mathcal{A}_F$ on $S \otimes \mathcal{H}_F$ is not unitarily equivalent to its composition with complex conjugation. This can be seen from the fact that, in this representation, the dimension of the space on which $\mathbb{C}$ acts by λ is larger than that of the space on which it acts by $\bar{\lambda}$. It then follows that the restriction of α_x to $\mathbb{C} \subset \mathcal{A}_F$ has to be the identity automorphism.

Similarly, the restriction of α_x to $M_3(\mathbb{C}) \subset \mathcal{A}_F$ is given by an inner automorphism of the form $f \mapsto v_x\, f\, v_x^*$, where $v_x \in \mathrm{SU}(3)$ is only determined modulo the center $Z_3 \sim \mu_3$ of $\mathrm{SU}(3)$. The restriction of α_x to $\mathbb{H} \subset \mathcal{A}_F$ is given by an inner automorphism of the form $f \mapsto q_x\, f\, q_x^*$, where $q_x \in \mathrm{SU}(2)$ is only determined modulo the center $Z_2 \sim \mu_2$ of $\mathrm{SU}(2)$.
Thus, upon passing to the finite abelian cover $\tilde{M}$ of M corresponding to the morphism $\pi_1(M) \to Z_2 \times Z_3 \sim \mu_6$, one gets a unitary element $u = (1, q, v) \in C^\infty(M, \mathrm{SU}(\mathcal{A}_F))$ such that $\alpha(f) = \mathrm{Ad}(u) f \mathrm{Ad}(u)^*$ for all $f \in C^\infty(M, \mathcal{A}_F)$. Replacing U by $U\,\mathrm{Ad}(u)^*$ one can then assume that U commutes with all $f \in C^\infty(M, \mathcal{A}_F)$. The commutation with J still holds, so that we obtain $U\,\mathrm{Ad}(u)^* \in \mathcal{C}$, with $\mathcal{C}$ the commutant of the algebra of operators on $\mathcal{H}$ generated by $\mathcal{A}$ and $J\mathcal{A}J^{-1}$. $\square$

15.2. Discrete part of the inner fluctuations and the Higgs field.

At a point $x \in M$ consider elements $a_i(x) = (\lambda_i, q_i, m_i)$ and $a_i'(x) = (\lambda_i', q_i', m_i')$ in $\mathcal{A}_F$. We introduce the notation

$$\varphi_1 = \sum \lambda_i(\alpha_i' - \lambda_i'), \quad \varphi_2 = \sum \lambda_i \beta_i' \tag{1.705}$$

$$\varphi_1' = \sum \alpha_i(\lambda_i' - \alpha_i') + \beta_i \bar{\beta}_i', \quad \varphi_2' = \sum(-\alpha_i \beta_i' + \beta_i(\bar{\lambda}_i' - \bar{\alpha}_i')), \tag{1.706}$$

where we used the notation (1.645) for quaternions.

PROPOSITION 1.200. *At a point $x \in M$ one obtains on the subspace corresponding to $\mathcal{H}_f \subset \mathcal{H}_F$ the expression*

$$\sum_i a_i[\gamma_5 \otimes D_F, a_i'](x)|_{\mathcal{H}_f} = \gamma_5 \otimes (A_q^{(0,1)} + A_\ell^{(0,1)}). \tag{1.707}$$

The term $A_q^{(0,1)}$ is defined on the $\mathcal{H}_{F,q}$ part and is of the form

$$A_q^{(0,1)} = \begin{pmatrix} 0 & X \\ X' & 0 \end{pmatrix} \otimes 1_3\,, \tag{1.708}$$

with

$$X = \begin{pmatrix} \Upsilon_u^* \varphi_1 & \Upsilon_u^* \varphi_2 \\ -\Upsilon_d^* \bar{\varphi}_2 & \Upsilon_d^* \bar{\varphi}_1 \end{pmatrix}, \quad \text{and} \quad X' = \begin{pmatrix} \Upsilon_u \varphi_1' & \Upsilon_d \varphi_2' \\ -\Upsilon_u \bar{\varphi}_2' & \Upsilon_d \bar{\varphi}_1' \end{pmatrix}. \tag{1.709}$$

The term $A_\ell^{(0,1)}$ is defined on the $\mathcal{H}_{F,\ell}$ part and is of the form

$$A_1^{(0,1)} = \begin{pmatrix} 0 & Y \\ Y' & 0 \end{pmatrix}, \tag{1.710}$$

with

$$Y = \begin{pmatrix} \Upsilon_\nu^* \varphi_1 & \Upsilon_\nu^* \varphi_2 \\ -\Upsilon_e^* \bar{\varphi}_2 & \Upsilon_e^* \bar{\varphi}_1 \end{pmatrix} \quad \text{and} \quad Y' = \begin{pmatrix} \Upsilon_\nu \varphi_1' & \Upsilon_e \varphi_2' \\ -\Upsilon_\nu \bar{\varphi}_2' & \Upsilon_e \bar{\varphi}_1' \end{pmatrix}, \tag{1.711}$$

with the φ as in (1.705) *and* (1.706).

PROOF. The result follows by a direct computation. $\square$

Notice that, according to Corollary 1.188 above, the off-diagonal part of D_F, which involves Υ_R, does not contribute to the inner fluctuations, since it exactly commutes with the algebra $\mathcal{A}_F$. Since the action of $\mathcal{A}_F$ on $\mathcal{H}_{\bar{f}}$ exactly commutes with D_F, it does not contribute to $A^{(0,1)}$. This means, as remarked after Corollary 1.188 that the presence of the Majorana mass term Υ_R does not generate new boson fields.

We now give the following parameterization of the discrete part $A^{(0,1)}$ of the inner fluctuations, which will be useful in the physical interpretation as Higgs bosons.

PROPOSITION 1.201. *Let $A^{(0,1)}$ be the discrete part of the inner fluctuations of the product geometry $M \times F$, as in Proposition 1.200 above.*

(1) *The inner fluctuations $A^{(0,1)}$ are parameterized by an arbitrary quaternion-valued function*

$$H \in C^\infty(M, \mathbb{H}), \qquad \text{with} \quad H = \varphi_1 + \varphi_2\, j\,, \quad \text{for} \quad \varphi_1, \varphi_2 \in C^\infty(M, \mathbb{C}). \tag{1.712}$$

(2) *The second line of the 2×2 matrix X in (1.709) is obtained from the first line upon replacing Υ_u with Υ_d and changing*

$$H \mapsto j\, H =: \tilde{H}. \tag{1.713}$$

PROOF. (1) First one checks that there are no linear relations between the four terms (1.705) and (1.706). We consider a single term $a[D_F, a']$. Taking $a = (\lambda, 0, 0)$ and $a' = (\lambda', 0, 0)$ gives

$$\varphi_1 = -\lambda\lambda' \quad \text{and} \quad \varphi_2 = \varphi_1' = \varphi_2' = 0.$$

Taking $a = (\lambda, 0, 0)$ and $a' = (0, j\,\bar{\beta}', 0)$ gives

$$\varphi_1' = \lambda\beta' \quad \text{and} \quad \varphi_1 = \varphi_1' = \varphi_2' = 0.$$

Similarly, taking $a = (0, \alpha, 0)$ and $a' = (\lambda', 0, 0)$ gives

$$\varphi_1' = \alpha\lambda' \quad \text{and} \quad \varphi_1 = \varphi_2 = \varphi_2' = 0.$$

Taking $a = (0, j\,\bar{\beta}, 0)$ and $a' = (\lambda', 0, 0)$ gives

$$\varphi_2' = \beta\bar{\lambda}' \quad \text{and} \quad \varphi_1 = \varphi_2 = \varphi_1' = 0.$$

This shows that the vector space $\Omega_D^{(0,1)}$ of linear combinations $\sum_i a_i[D_F, a_i']$ is the space of pairs of quaternion-valued functions $q(x)$ and $q'(x)$.
The self-adjoint condition $A = A^*$ is equivalent to $q' = q^*$ and we see that the discrete part $A^{(0,1)}$ is exactly given by a single quaternion-valued function $H(x) \in \mathbb{H}$ for $x \in M$.
(2) The transition is given by $(\varphi_1, \varphi_2) \mapsto (-\bar{\varphi}_2, \bar{\varphi}_1)$, which corresponds to the multiplication of $H = \varphi_1 + \varphi_2\, j$ by j on the left. ☐

REMARK 1.202. The property in (2) of Proposition 1.201 shows that the role of H in the coupling of the up-part is related to its role in the coupling of the down-part by the replacement $H \mapsto jH$. From the physics viewpoint this shows that one needs to change the hypercharge of the Higgs doublet to its opposite in order to couple it to the up quark, as is known in the extension of the Glashow–Weinberg–Salam model from leptons to quarks.

REMARK 1.203. In particular, the result of Proposition 1.200 shows that, in the case of a Dirac operator $D = D(\Upsilon)$ where the term $\Upsilon_R = 0$, the operator D is equivalent to $D \equiv 0$ up to inner fluctuations. This shows an example of the general fact mentioned in Remark 1.143.

Notice that the notation we use here differs from that of [**45**]. The Higgs doublet denoted by H in [**45**] is in our notation the doublet $\tilde{H}$ of (1.713) above. This is especially relevant in the calculations of §16.1 below.

15.3. Powers of $D^{(0,1)}$.

For later purposes, we compute here the trace of powers of the perturbed Dirac operator

$$D^{(0,1)} := D + A^{(0,1)} + JA^{(0,1)}J. \tag{1.714}$$

PROPOSITION 1.204. *The traces of the second and fourth powers of $D^{(0,1)}$ are given by the formulae*

$$\mathrm{Tr}((D^{(0,1)})^2) = 4\,a\,|1+H|^2 + 2\,c \tag{1.715}$$

and

$$\mathrm{Tr}((D^{(0,1)})^4) = 4\,b\,|1+H|^4 \, + 2\,d + 8\,e\,|1+H|^2\,, \tag{1.716}$$

where H is as in Proposition 1.201 and the coefficients a, b, c, d, e are of the form

$$\begin{aligned} a =\ & \mathrm{Tr}(\Upsilon_\nu^*\Upsilon_\nu + \Upsilon_e^*\Upsilon_e + 3(\Upsilon_u^*\Upsilon_u + \Upsilon_d^*\Upsilon_d)) \\ b =\ & \mathrm{Tr}((\Upsilon_\nu^*\Upsilon_\nu)^2 + (\Upsilon_e^*\Upsilon_e)^2 + 3(\Upsilon_u^*\Upsilon_u)^2 + 3(\Upsilon_d^*\Upsilon_d)^2) \\ c =\ & \mathrm{Tr}(\Upsilon_R^*\Upsilon_R) \\ d =\ & \mathrm{Tr}((\Upsilon_R^*\Upsilon_R)^2) \\ e =\ & \mathrm{Tr}(\Upsilon_R^*\Upsilon_R\Upsilon_\nu^*\Upsilon_\nu). \end{aligned} \tag{1.717}$$

PROOF. We first show that for the restriction $D_q^{(0,1)}$ to the subspace $\mathcal{H}_{F,q} \subset \mathcal{H}_F$ we have

$$\mathrm{Tr}((D_q^{(0,1)})^2) =\ 12\,|1+H|^2\,\mathrm{Tr}(\Upsilon_u^*\Upsilon_u + \Upsilon_d^*\Upsilon_d) \tag{1.718}$$

and

$$\mathrm{Tr}((D_q^{(0,1)})^4) =\ 12\,|1+H|^4\,\mathrm{Tr}((\Upsilon_u^*\Upsilon_u)^2 + (\Upsilon_d^*\Upsilon_d)^2). \tag{1.719}$$

Notice that the left-hand side of (1.718) is obtained from $2\,\mathrm{Tr}((A_q^{(0,1)})^2)$, after replacing H by $1+H$ to take into account the presence of the operator $D_q = D|_{\mathcal{H}_{F,q}}$. The product XX^*, for X as in (1.709), is given by the diagonal matrix

$$\begin{aligned} XX^* =\ & \begin{pmatrix} \Upsilon_u^*\Upsilon_u\,(\varphi_1\bar\varphi_1 + \varphi_2\bar\varphi_2) & 0 \\ 0 & \Upsilon_d^*\Upsilon_d\,(\varphi_1\bar\varphi_1 + \varphi_2\bar\varphi_2) \end{pmatrix} \\ =\ & |H|^2 \begin{pmatrix} \Upsilon_u^*\Upsilon_u & 0 \\ 0 & \Upsilon_d^*\Upsilon_d \end{pmatrix}. \end{aligned} \tag{1.720}$$

One has $\mathrm{Tr}((A_q^{(0,1)})^2) = 3\,\mathrm{Tr}(XX^* + X^*X) = 6\,\mathrm{Tr}(XX^*)$. This gives the equality (1.718). Similarly, one has $\mathrm{Tr}((A_3^{(0,1)})^4) = 3\,\mathrm{Tr}((XX^*)^2 + (X^*X)^2) = 6\,\mathrm{Tr}((XX^*)^2)$, which gives the identity (1.719).

We then show that, for the restriction $D_\ell^{(0,1)}$ of $D^{(0,1)}$ to the subspace $\mathcal{H}_{F,\ell} \subset \mathcal{H}_F$ we have

$$\mathrm{Tr}((D_\ell^{(0,1)})^2) =\ 4\,|1+H|^2\,\mathrm{Tr}(\Upsilon_\nu^*\Upsilon_\nu + \Upsilon_e^*\Upsilon_e) + 2\,\mathrm{Tr}(\Upsilon_R^*\Upsilon_R), \tag{1.721}$$

and

$$\begin{aligned}\operatorname{Tr}((D_\ell^{(0,1)})^4) = \;& 4\,|1+H|^4 \operatorname{Tr}((\Upsilon_\nu^*\Upsilon_\nu)^2 + (\Upsilon_e^*\Upsilon_e)^2) + 2\operatorname{Tr}((\Upsilon_R^*\Upsilon_R)^2) \\ & +8\,|1+H|^2 \operatorname{Tr}(\Upsilon_R^*\Upsilon_R\Upsilon_\nu^*\Upsilon_\nu).\end{aligned} \tag{1.722}$$

Consider the decomposition of $\mathcal{H}_{F,\ell}$ into the eight subspaces of the form

$$\begin{array}{ccccccccc} & (\,|\uparrow\rangle_R\otimes\mathbf{1}^0 & \oplus & |\downarrow\rangle_R\otimes\mathbf{1}^0 & \oplus & |\uparrow\rangle_L\otimes\mathbf{1}^0 & \oplus & |\downarrow\rangle_L\otimes\mathbf{1}^0 & \\ \oplus & \mathbf{1}\otimes|\uparrow\rangle_R & \oplus & \mathbf{1}\otimes|\downarrow\rangle_R & \oplus & \mathbf{1}\otimes|\uparrow\rangle_L & \oplus & \mathbf{1}\otimes|\downarrow\rangle_L\,)\otimes\mathbb{C}^3. \end{array}$$

In this decomposition, we can write the operator $D_\ell^{(0,1)}$ as the matrix

$$\begin{pmatrix} 0 & 0 & \Upsilon_\nu^*\varphi_1 & \Upsilon_\nu^*\varphi_2 & \Upsilon_R^* & 0 & 0 & 0 \\ 0 & 0 & -\Upsilon_e^*\bar\varphi_2 & \Upsilon_e^*\bar\varphi_1 & 0 & 0 & 0 & 0 \\ \Upsilon_\nu\bar\varphi_1 & -\Upsilon_e\varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Upsilon_\nu\bar\varphi_2 & \Upsilon_e\varphi_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Upsilon_R & 0 & 0 & 0 & 0 & 0 & \bar\Upsilon_\nu^*\bar\varphi_1 & \bar\Upsilon_\nu^*\bar\varphi_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\bar\Upsilon_e^*\varphi_2 & \bar\Upsilon_e^*\varphi_1 \\ 0 & 0 & 0 & 0 & \bar\Upsilon_\nu\varphi_1 & -\bar\Upsilon_e\bar\varphi_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bar\Upsilon_\nu\varphi_2 & \bar\Upsilon_e\bar\varphi_1 & 0 & 0 \end{pmatrix},$$

where $\bar\Upsilon$ denotes the complex conjugate matrix.
The only matrix elements of the square of $D_\ell^{(0,1)}$ involving Υ_R or Υ_R^* are collected in the matrix

$$\begin{pmatrix} S_{11} & 0 & 0 & 0 & 0 & 0 & \Upsilon_R^*\bar\Upsilon_\nu^*\bar\varphi_1 & \Upsilon_R^*\bar\Upsilon_\nu^*\bar\varphi_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Upsilon_\nu\Upsilon_R^*\bar\varphi_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Upsilon_\nu\Upsilon_R^*\bar\varphi_2 & 0 & 0 & 0 \\ 0 & 0 & \Upsilon_R\Upsilon_\nu^*\varphi_1 & \Upsilon_R\Upsilon_\nu^*\varphi_2 & S_{55} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \bar\Upsilon_\nu\Upsilon_R\varphi_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \bar\Upsilon_\nu\Upsilon_R\varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where

$$S_{11} = \Upsilon_R^*\Upsilon_R + \Upsilon_\nu^*\Upsilon_\nu|H|^2\,,\quad S_{55} = \bar\Upsilon_\nu^*\bar\Upsilon_\nu|H|^2 + \Upsilon_R\Upsilon_R^*\,.$$

This shows that, for $\operatorname{Tr}((D_\ell^{(0,1)})^2)$, one only gets two additional terms involving Υ_R and each of them gives $\operatorname{Tr}(\Upsilon_R\Upsilon_R^*)$. The trace $\operatorname{Tr}((D_\ell^{(0,1)})^4)$ is the Hilbert-Schmidt norm squared of $(D_\ell^{(0,1)})^2$. We just need to add to the terms coming from the same computation in (1.718) the contribution of the terms involving Υ_R. The term S_{11} contributes (after replacing $H \mapsto 1+H$) by $2|1+H|^2\operatorname{Tr}(\Upsilon_R^*\Upsilon_R\Upsilon_\nu^*\Upsilon_\nu)$ and $\operatorname{Tr}((\Upsilon_R^*\Upsilon_R)^2)$. The term S_{55} gives a similar contribution. All the other terms give simple additive contributions. One gets the result using

$$\operatorname{Tr}(\bar\Upsilon_\nu\Upsilon_R\Upsilon_R^*\bar\Upsilon_\nu^*) = \operatorname{Tr}(\Upsilon_R^*\Upsilon_R\Upsilon_\nu^*\Upsilon_\nu),$$

which follows, using complex conjugation, from the fact that Υ_R is a symmetric matrix, so that $\bar\Upsilon_R = \Upsilon_R^*$. □

15.4. Continuous part of the inner fluctuations and gauge bosons.

Let us now consider the $A^{(1,0)}$ part of the inner fluctuations of the product geometry $M \times F$. We show that it gives the gauge bosons of the Standard Model coupled to the correct values of the hypercharges.

PROPOSITION 1.205. *Consider inner fluctuations of the form*

$$A^{(1,0)} = \sum_i a_i [\partial\!\!\!/_M \otimes 1, a_i'], \tag{1.723}$$

with $a_i = (\lambda_i, q_i, m_i)$, $a_i' = (\lambda_i', q_i', m_i')$ *elements of* $\mathcal{A} = C^\infty(M, \mathcal{A}_F)$. *The expression* (1.723) *determines the following data.*

(1) *A* U(1) *gauge field*

$$\Lambda = \sum_i \lambda_i \, d\lambda_i' \tag{1.724}$$

(2) *An* SU(2) *gauge field*

$$Q = \sum_i q_i \, dq_i' \tag{1.725}$$

(3) *A* U(3) *gauge field*

$$V' = \sum_i m_i \, dm_i'. \tag{1.726}$$

PROOF. For (1) notice that we have two expressions to compute, since there are two different actions of $\lambda(x)$ on $L^2(M, S)$, given respectively by

$$\xi(x) \mapsto \lambda(x)\,\xi(x) \quad \text{and} \quad \xi(x) \mapsto \bar{\lambda}(x)\,\xi(x). \tag{1.727}$$

For the first action in (1.727), one finds using (1.580), that the expression

$$\Lambda = \sum_i \lambda_i [\partial\!\!\!/_M \otimes 1, \lambda_i']$$

is of the form

$$\Lambda = \sqrt{-1} \sum_i \lambda_i \partial_\mu \lambda_i' \gamma^\mu = \Lambda_\mu \gamma^\mu.$$

It is self-adjoint when the scalar functions

$$\Lambda_\mu = \sqrt{-1} \sum_i \lambda_i \partial_\mu \lambda_i'$$

are real-valued.
It follows then that the second action in (1.727) is given by

$$\sum_i \bar{\lambda}_i [\partial\!\!\!/_M \otimes 1, \bar{\lambda}_i'] = \sqrt{-1} \sum_i \bar{\lambda}_i \partial_\mu \bar{\lambda}_i' \gamma^\mu = -\Lambda_\mu \gamma^\mu.$$

Thus, we see that, even though we have two representations of the $\lambda(x)$, these generate only one U(1) gauge potential.
For (2) notice that the action of quaternions $\mathbb{H}$ can be represented in the form

$$q = f_0 + \sum_\alpha i f_\alpha \sigma^\alpha, \quad \text{with} \quad f_0, f_\alpha \in C^\infty(M, \mathbb{R}), \tag{1.728}$$

where σ^α are the Pauli matrices (1.628) and $i = \sqrt{-1}$.

The Pauli matrices are self-adjoint, hence the terms of the form

$$f_0[\partial\!\!\!/_M \otimes 1, i f'_\alpha \sigma^\alpha]$$

are also self-adjoint.

The elements $\{1, i\sigma^\alpha\}$ form a basis for the algebra $\mathbb{H}$ of quaternions. Thus, since the elements of this basis commute with $\partial\!\!\!/_M$, one can rewrite

$$\sum_i q_i[\partial\!\!\!/_M \otimes 1, q'_i] = f_0[\partial\!\!\!/_M \otimes 1, f'_0] + \sum_\alpha f_\alpha[\partial\!\!\!/_M \otimes 1, i f'_\alpha \sigma^\alpha],$$

where all the f and f' are real-valued functions. Thus, the self-adjoint part of this expression is given by

$$Q = \sum_\alpha f_\alpha[\partial\!\!\!/_M \otimes 1, i f'_\alpha \sigma^\alpha],$$

which is an SU(2) gauge field.

For (3), the result follows as a special case of the computation of expressions of the form

$$A = \sum_i a_i[\partial\!\!\!/_M \otimes 1, a'_i], \quad \text{with} \quad a_i, a'_i \in C^\infty(M, M_N(\mathbb{C})).$$

One obtains Clifford multiplication by all matrix-valued 1-forms on M in this manner. The self-adjointness condition $A = A^*$ then reduces them to taking values in the Lie algebra of $U(N)$, by setting $A = iT$ and using the identification

$$\operatorname{Lie} U(N) = \{T \in M_N(\mathbb{C}) \mid T^* = -T\}.$$

□

We introduce the notation

$$\Lambda_\mu = \frac{g_1}{2} B_\mu \tag{1.729}$$

for the U(1) gauge potential of Proposition 1.205. Physically it gives the generator of hypercharge, not to be confused with the electromagnetic vector potential. Similarly, we denote the SU(2) gauge field of Proposition 1.205 by the notation

$$Q = Q_\mu \gamma^\mu, \quad \text{where} \quad Q_\mu = \frac{g_2}{2} W_\mu^\alpha \sigma^\alpha. \tag{1.730}$$

Using (1.580), we see that this corresponds to the covariant derivative

$$\partial_\mu - \frac{i}{2} g_2 W_\mu^\alpha \sigma^\alpha. \tag{1.731}$$

We now explain how to reduce the gauge field V' of (1.726) to the Lie subalgebra $\operatorname{Lie}\mathrm{SU}(3)$ of $\operatorname{Lie}\mathrm{U}(3)$. We consider the following analogue of Definition 1.184 of the unimodular subgroup $\mathrm{SU}(\mathcal{A}_F)$.

DEFINITION 1.206. *A gauge potential A is "unimodular" if* $\mathrm{Tr}(A) = 0$.

We can now parameterize the unimodular gauge potentials and their adjoint action, i.e. the combinations of the form $A + JAJ^{-1}$.

PROPOSITION 1.207. (1) *The unimodular gauge potentials are parameterized by a* U(1) *gauge field B, an* SU(2) *gauge field W and an* SU(3) *gauge field V.*

(2) *The adjoint action* $A+JAJ^{-1}$ *on* $\mathcal{H}_f$ *is obtained by replacing* ∂_μ *by* $\partial_\mu+\mathbb{A}_\mu$ *where*

$$\mathbb{A}_\mu = (\mathbb{A}_\mu^q \oplus \mathbb{A}_\mu^\ell) \otimes 1_3, \tag{1.732}$$

where the 1_3 *is for the* $N=3$ *generations, and* $\mathbb{A}_\mu^q$ *and* $\mathbb{A}_\mu^\ell$ *are given respectively by*

(1.733)

$$\mathbb{A}_\mu^q = \begin{pmatrix} -\frac{2i}{3}g_1 B_\mu \otimes 1_3 & 0 & 0 \\ 0 & \frac{i}{3}g_1 B_\mu \otimes 1_3 & 0 \\ 0 & 0 & (-\frac{i}{2}g_2 W_\mu^\alpha \sigma^\alpha - \frac{i}{6}g_1 B_\mu \otimes 1_2) \otimes 1_3 \end{pmatrix}$$
$$+ \quad 1_4 \otimes \left(-\tfrac{i}{2}g_3 V_\mu^i \lambda^i\right),$$

$$\mathbb{A}_\mu^\ell = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i g_1 B_\mu & 0 \\ 0 & 0 & (-\frac{i}{2}g_2 W_\mu^\alpha \sigma^\alpha + \frac{i}{2}g_1 B_\mu \otimes 1_2) \end{pmatrix}, \tag{1.734}$$

where in (1.733) *and* (1.734) *the* σ^α *are the Pauli matrices* (1.628) *and the* λ^i *are the Gell-Mann matrices* (1.421).

PROOF. 1) The action of A on the subspace $\mathcal{H}_f$ is of the form

$$\begin{pmatrix} \Lambda & 0 & 0 & 0 \\ 0 & -\Lambda & 0 & 0 \\ 0 & 0 & Q_{11} & Q_{12} \\ 0 & 0 & Q_{21} & Q_{22} \end{pmatrix} \tag{1.735}$$

on leptons and quarks, with Λ and Q as in Proposition 1.205, cf. (1.729) and (1.730) above.
Thus, (1.735) is traceless, since Q is traceless as a linear combination of the Pauli matrices. The action of A on the subspace $\mathcal{H}_{\bar f}$ is given by Λ on the subspace of leptons and by V' on the space of quarks. One has 4 leptons and 4 quarks per generation because of the two possible chiralities, while the color index is taken care of by V'. Thus, the unimodularity condition means that we have

$$3 \cdot 4 \cdot (\Lambda + \mathrm{Tr}(V')) = 0.$$

Thus, we can write V' as a sum of the form

$$V' = -V - \frac{1}{3}\begin{pmatrix} \Lambda & 0 & 0 \\ 0 & \Lambda & 0 \\ 0 & 0 & \Lambda \end{pmatrix} = -V - \frac{1}{3}\Lambda 1_3, \tag{1.736}$$

where V is traceless, i.e. it is an SU(3) gauge potential.
2) Since the charge conjugation antilinear operator J_M commutes with $\partial\!\!\!/_M$, it anticommutes with the γ_μ and the conjugation by J introduces an additional minus sign in the gauge potentials. The computation of $A+JAJ^{-1}$ gives, on quarks and leptons respectively, the matrices

$$\begin{pmatrix} \Lambda - V' & 0 & 0 & 0 \\ 0 & -\Lambda - V' & 0 & 0 \\ 0 & 0 & Q_{11} - V' & Q_{12} \\ 0 & 0 & Q_{21} & Q_{22} - V' \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -2\Lambda & 0 & 0 \\ 0 & 0 & Q_{11}-\Lambda & Q_{12} \\ 0 & 0 & Q_{21} & Q_{22}-\Lambda \end{pmatrix}.$$

Thus, using (1.736), we obtain for the $(1,0)$-part of the inner fluctuation $A+JAJ^{-1}$ of the metric the matrices

$$\begin{pmatrix} \frac{4}{3}\Lambda+V & 0 & 0 & 0 \\ 0 & -\frac{2}{3}\Lambda+V & 0 & 0 \\ 0 & 0 & Q_{11}+\frac{1}{3}\Lambda+V & Q_{12} \\ 0 & 0 & Q_{21} & Q_{22}+\frac{1}{3}\Lambda+V \end{pmatrix}, \tag{1.737}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -2\Lambda & 0 & 0 \\ 0 & 0 & Q_{11}-\Lambda & Q_{12} \\ 0 & 0 & Q_{21} & Q_{22}-\Lambda \end{pmatrix}. \tag{1.738}$$

This completes the proof. □

Notice that one can see from the entries of the matrices (1.737) and (1.738) that the gauge bosons are coupled to the correct hypercharges. In fact, we recover the correct values as in (1.659) (cf. also (1.415)).

15.5. Independence of the boson fields.

It remains to show that the fields $H=\varphi_1+\varphi_2\, j$ of Proposition 1.201 and B,W,V of Proposition 1.207 are independent.

Proposition 1.208. *The unimodular inner fluctuations of the metric are parameterized by independent fields φ_1, φ_2, B, W, V, as in Propositions 1.201 and 1.207.*

Proof. Let Z be the real vector bundle over M, with fiber at x

$$\mathbb{C}\oplus\mathbb{C}\oplus T_x^*M\oplus T_x^*M\otimes \operatorname{Lie}\mathrm{SU}(2)\oplus T_x^*M\otimes\operatorname{Lie}\mathrm{SU}(3).$$

By construction the inner fluctuations are sections of the bundle Z.
The space of sections $\mathcal{S}$ obtained from inner fluctuations is in fact not just a linear space over $\mathbb{R}$, but also a module over the algebra $C^\infty(M,\mathbb{R})$ which is the real part of $C^\infty(M,\mathcal{A}_F)$ as in Lemma 1.198. Indeed, the inner fluctuations are obtained as expressions of the form $A=\sum_i a_i\,[D,a_i']$. One has to check that left multiplication by $f\in C^\infty(M,\mathbb{R})$ does not alter the self-adjointness condition $A=A^*$. This follows from Proposition 1.125, since we are replacing a_j by fa_j, where f commutes with A and is real so that $f=f^*$.
To show that $\mathcal{S}=C^\infty(M,Z)$ it is enough to know that one can find sections in $\mathcal{S}$ that span the full vector space Z_x at any given point $x\in M$. Then $C^\infty(M,\mathbb{R})$-linearity shows that the same sections continue to span the nearby fibers. Using a partition of unity one can then express any global section of Z as an element of $\mathcal{S}$.
Choose first the elements $a_i(y)=(\lambda_i,q_i,m_i)$, $a_i'(y)=(\lambda_i',q_i',m_i')$ independent of $y\in N(x)$ in a neighborhood of x. Using Proposition 1.201, one knows that $H(x)$ can be an arbitrary element of $\mathbb{H}$, while $B(x),W(x),V(x)$ all vanish because they are differential expressions in the a_i'.

The independence of λ, q and m in the formulae (1.724), (1.725), (1.726) implies that one can construct arbitrary $B(x)$, $W(x)$, $V(x)$ in the form $\sum_i a_i[D, a_i']$. These, however, will not suffice to give an arbitary value for φ_1 and φ_2, but this can be corrected by adding an element of the form described above, with vanishing B, W, and V. $\square$

15.6. The Dirac operator and its square.

We now consider the full (perturbed) Dirac operator D_A, where the original Dirac operator $D = \partial\!\!\!/_M \otimes 1 + \gamma_5 \otimes D_F$ of the product geometry is twisted by an inner fluctuation. The operator D_A is of the form

$$D_A = D^{(1,0)} + \gamma_5 \otimes D^{(0,1)} \tag{1.739}$$

where $D^{(0,1)}$ is given by (1.714) and $D^{(1,0)}$ is of the form,

$$D^{(1,0)} = \sqrt{-1}\gamma^\mu(\nabla^s_\mu + \mathbb{A}_\mu), \tag{1.740}$$

where ∇^s is the spin connection (cf. (1.580)). The gauge potential $\mathbb{A}_\mu$ splits as a direct sum in the decomposition associated to $\mathcal{H}_F = \mathcal{H}_f \oplus \mathcal{H}_{\bar f}$ and its restriction to $\mathcal{H}_f$ is given by the expression as in Proposition 1.207.
We also consider the operator D_A^2. To write the components in a convenient notation, we introduce the following.

Definition 1.209. *For $\varphi = (\varphi_1, \varphi_2)$ as in §15.2 above and M_1, M_2 a pair of 3×3 matrices, we set*

$$T(M_1, M_2, \varphi) = \begin{pmatrix} 0 & 0 & M_1^*\varphi_1 & M_1^*\varphi_2 \\ 0 & 0 & -M_2^*\bar\varphi_2 & M_2^*\bar\varphi_1 \\ M_1\bar\varphi_1 & -M_2\varphi_2 & 0 & 0 \\ M_1\bar\varphi_2 & M_2\varphi_1 & 0 & 0 \end{pmatrix}. \tag{1.741}$$

We also define

$$\begin{aligned} \mathbb{M}(\varphi) &= T(\Upsilon_u, \Upsilon_d, \varphi) \otimes 1_3 \oplus T(\Upsilon_\nu, \Upsilon_e, \varphi) \\ &\oplus \overline{T(\Upsilon_u, \Upsilon_d, \varphi)} \otimes 1_3 \oplus \overline{T(\Upsilon_\nu, \Upsilon_e, \varphi)}. \end{aligned} \tag{1.742}$$

By construction, $\mathbb{M}(\varphi)$ is self-adjoint and one has

$$\mathrm{Tr}(\mathbb{M}(\varphi)^2) = 4\,a\,|\varphi|^2\,, \tag{1.743}$$

where the coefficient a is as in (1.717).
We then have the following result that gives the explicit form of the operator D_A^2.

Lemma 1.210. *The square of D_A is given by*

$$D_A^2 = \nabla^*\nabla - E, \tag{1.744}$$

where $\nabla^*\nabla$ is the connection Laplacian for the connection

$$\nabla = \nabla^s + \mathbb{A} \tag{1.745}$$

and the endomorphism E is given by

$$-E = \frac{1}{4}\, s \otimes \mathrm{id} + \sum_{\mu<\nu} \gamma^\mu\gamma^\nu \otimes \mathbb{F}_{\mu\nu} - i\,\gamma_5\,\gamma^\mu \otimes \mathbb{M}(D_\mu\varphi) + 1_4 \otimes (D^{0,1})^2, \tag{1.746}$$

where $s = -R$ is the scalar curvature and $H = \varphi_1 + \varphi_2 j$ as above, with $\varphi = (\varphi_1, \varphi_2)$. Here $\mathbb{F}_{\mu\nu}$ is the curvature of the connection $\mathbb{A}$ and $\varphi = (\varphi_1, \varphi_2)$ is a row vector. The term $D_\mu \varphi$ in (1.746) is of the form

$$D_\mu \varphi = \partial_\mu \varphi + \frac{i}{2} g_2 W_\mu^\alpha \varphi \, \sigma^\alpha - \frac{i}{2} g_1 B_\mu \, \varphi. \tag{1.747}$$

PROOF. By construction $D^{1,0}$ anticommutes with γ_5. Thus, one has

$$D_A^2 = (D^{1,0})^2 + 1_4 \otimes (D^{0,1})^2 - \gamma_5 \, [D^{1,0}, 1_4 \otimes D^{0,1}].$$

The last term is of the form

$$[D^{1,0}, 1_4 \otimes D^{0,1}] = \sqrt{-1} \, \gamma^\mu \, [(\nabla_\mu^s + \mathbb{A}_\mu), 1_4 \otimes D^{0,1}].$$

Using (1.583), one can replace ∇_μ^s by ∂_μ without changing the result. In order to compute the commutator $[\mathbb{A}_\mu, D^{0,1}]$, notice first that the off-diagonal term of $D^{0,1}$ does not contribute, since the corresponding matrix elements of $\mathbb{A}_\mu^\ell$ are zero. Thus, it is enough to compute the commutator of the following matrix $\mathbb{W}$:

$$\begin{pmatrix} -\frac{i}{2} g_1 B_\mu & 0 & 0 & 0 \\ 0 & \frac{i}{2} g_1 B_\mu & 0 & 0 \\ 0 & 0 & -\frac{i}{2} g_2 W_\mu^3 & -\frac{i}{2} g_2 (W_\mu^1 - i W_\mu^2) \\ 0 & 0 & -\frac{i}{2} g_2 (W_\mu^1 + i W_\mu^2) & \frac{i}{2} g_2 W_\mu^3 \end{pmatrix} \tag{1.748}$$

with a matrix of the form $T(M_1, M_2, \varphi)$ as in (1.741). One gets

$$[\mathbb{W}, T(M_1, M_2, \varphi)] = T(M_1, M_2, \chi), \tag{1.749}$$

where

$$(\chi_1, \chi_2) = -\frac{i}{2} g_1 B_\mu (\varphi_1, \varphi_2) + \frac{i}{2} g_2 W_\mu^\alpha \, (\varphi_1, \varphi_2) \, \sigma^\alpha. \tag{1.750}$$

□

16. The spectral action and the Standard Model Lagrangian

In this section we compute explicitly the expansion of Theorem 1.145 for the spectral action for the product geometry $M \times F$. We then show how to modify the spectral action functional to include the fermionic action and how this can be done in such a way as to solve the fermion doubling problem and obtain the correct counting of the fermionic degrees of freedom through the use of Pfaffians as a replacement for the formulation of Majorana fermions in the Euclidean setting. As in the previous sections, we follow closely [**52**].

16.1. The asymptotic expansion of the spectral action on $M \times F$.

The main result of this section is Theorem 1.212 below which computes the explicit form of Theorem 1.145 in the case of the spectral action associated to the product geometry $M \times F$.

We first recall the notation we have been using in the previous sections. We let f_k with $k = 0, \ldots, 4$ be the momenta (1.533) of the test function f in the spectral action functional (1.526). As in §9.8, we let $C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma}$ be the square of the Weyl curvature tensor and we let $R^* R^*$ be the form (1.599) that integrates to the Euler characteristic as in (1.600). We let $G_{\mu\nu}^i$, $F_{\mu\nu}^\alpha$, $B_{\mu\nu}$ denote, respectively, the curvatures associated to the SU(3), SU(2) and U(1) gauge fields of §15.4. Let $\{a,b,c,d,e\}$ be the coefficients as in (1.717) and let $D_\mu\varphi$ be the expression defined in (1.747).

We prove a preliminary lemma, which will be useful in applying Gilkey's theorem (cf. Theorem 1.154 above) to the proof of Theorem 1.212 below.

LEMMA 1.211. *Consider the expression*

$$E' = -\tfrac{R}{6}\mathrm{id} + E = \tag{1.751}$$
$$(\tfrac{R}{12}\mathrm{id} - 1_4 \otimes (D^{0,1})^2) - \textstyle\sum_{\mu<\nu} \gamma^\mu\gamma^\nu \otimes \mathbb{F}_{\mu\nu} + i\,\gamma_5\,\gamma^\mu \otimes \mathbb{M}(D_\mu\,\varphi)$$

and let

$$\Sigma = \frac{f_2}{8\,\pi^2}\Lambda^2\mathrm{Tr}(E') + \frac{f_0}{32\,\pi^2}\mathrm{Tr}((E')^2). \tag{1.752}$$

One obtains

$$\begin{aligned}\Sigma \;&=\; \frac{4\,f_2}{\pi^2}\,\Lambda^2\,R - \frac{f_2}{2\,\pi^2}\,\Lambda^2\,\mathrm{Tr}((D^{0,1})^2) + \frac{f_0}{8\,\pi^2}\,\mathrm{Tr}(\mathbb{M}(D_\mu\,\varphi)^2)\\ &+\; \frac{f_0}{8\,\pi^2}\,\mathrm{Tr}((\frac{R}{12} - (D^{0,1})^2)^2) + \frac{f_0}{16\,\pi^2}\,\mathrm{Tr}(\mathbb{F}_{\mu\nu}\overline{\mathbb{F}}^{\mu\nu}).\end{aligned} \tag{1.753}$$

PROOF. The contribution of $\mathrm{Tr}(E')$ is only coming from the first term of (1.751), since the trace of the two others vanishes due to the Clifford algebra terms. The coefficient of $\frac{f_2\,\Lambda^2}{\pi^2}\,R$ is equal to $\frac{1}{8} \times \frac{1}{12} \times 4 \times 96 = 4$. To get the contribution of $\mathrm{Tr}((E')^2)$, notice that the three terms of the sum (1.751) are pairwise orthogonal in the Clifford algebra, so that the trace of the square is just the sum of the three contributions from each of these terms. Again the factor of 4 comes from the dimension of spinors and the summation on all indices $\mu\nu$ gives a factor of two in the denominator for $\frac{f_0}{16\,\pi^2}$. ∎

We then obtain the following result.

THEOREM 1.212. *The expansion of Theorem 1.145 applied to the spectral action for the product geometry $M \times F$ gives a functional S of the form*

$$\begin{aligned}S \;&=\; \frac{1}{\pi^2}(48\,f_4\,\Lambda^4 - f_2\,\Lambda^2\,c + \frac{f_0}{4}\,d)\int \sqrt{g}\,d^4x\\ &+ \frac{96\,f_2\,\Lambda^2 - f_0\,c}{24\pi^2}\int R\,\sqrt{g}\,d^4x\\ &+\; \frac{f_0}{10\,\pi^2}\int(\frac{11}{6}\,R^*R^* - 3\,C_{\mu\nu\rho\sigma}\,C^{\mu\nu\rho\sigma})\,\sqrt{g}\,d^4x\\ &+\; \frac{(-2\,a\,f_2\,\Lambda^2\,+\,e\,f_0)}{\pi^2}\int |\varphi|^2\,\sqrt{g}\,d^4x\\ &+\; \frac{f_0}{2\,\pi^2}\int a\,|D_\mu\varphi|^2\,\sqrt{g}\,d^4x\\ &-\; \frac{f_0}{12\,\pi^2}\int a\,R\,|\varphi|^2\,\sqrt{g}\,d^4x\\ &+\; \frac{f_0}{2\,\pi^2}\int(g_3^2\,G^i_{\mu\nu}\,\overline{G}^{\mu\nu i} + g_2^2\,F^\alpha_{\mu\nu}\,\overline{F}^{\mu\nu\alpha} + \frac{5}{3}\,g_1^2\,B_{\mu\nu}\,\overline{B}^{\mu\nu})\,\sqrt{g}\,d^4x\\ &+\; \frac{f_0}{2\,\pi^2}\int b\,|\varphi|^4\,\sqrt{g}\,d^4x.\end{aligned} \tag{1.754}$$

PROOF. We prove the result using (1.744) and Gilkey's theorem 1.154 to compute the spectral action. By remark 1.155 the relevant term is $-\frac{1}{6}R + \mathcal{E}$ which is the sum (1.751) and we need to compute the sum (1.752). By Lemma 1.211 above, we know that this is of the form (1.753). Notice also that the curvature $\Omega_{\mu\nu}$ of the connection ∇ is independent of the additional term $D^{(0,1)}$.

We subdivide the proof into checking the different terms that arise from the asymptotic formula of Theorem 1.145 applied to the spectral action of the product geometry:

- Cosmological terms
- Riemannian curvature terms
- Higgs minimal coupling
- Higgs mass terms
- Higgs quartic potential
- Yang–Mills terms

Cosmological terms. The presence of the additional off-diagonal term in the Dirac operator of the finite geometry adds two contributions to the cosmological term of §11.4. Thus, while the dimension $N = 96$ contributes by the term

$$\frac{48}{\pi^2} f_4 \Lambda^4 \int_M \sqrt{g}\, d^4x\,,$$

we get the additional coefficients

$$-\frac{f_2}{\pi^2}\Lambda^2 \operatorname{Tr}(\Upsilon_R^* \Upsilon_R) = -\frac{c\, f_2}{\pi^2}\Lambda^2,$$

which are obtained from the second term of (1.753) using (1.715). Finally, we also get

$$\frac{f_0}{4\,\pi^2}\operatorname{Tr}((\Upsilon_R^*\Upsilon_R)^2) = \frac{d\, f_0}{4\,\pi^2}\,,$$

which comes from the fifth term in (1.753). Thus, the cosmological term gives

$$\frac{1}{\pi^2}(48\, f_4\,\Lambda^4 - f_2\,\Lambda^2\, \operatorname{Tr}(\Upsilon_R^*\Upsilon_R) + \frac{f_0}{4}\operatorname{Tr}((\Upsilon_R^*\Upsilon_R)^2)) \int_M \sqrt{g}\, d^4x\,. \tag{1.755}$$

Riemann curvature terms. we obtain the terms involving the Riemann curvature tensor from the spectral action as in §11.4. Thus, one has in this way a contribution of the form

$$\frac{1}{\pi^2}\int_M (4\, f_2\,\Lambda^2\, R - \frac{3}{10}\, f_0\, C_{\mu\nu\rho\sigma}\, C^{\mu\nu\rho\sigma})\sqrt{g}\, d^4x\,. \tag{1.756}$$

One also has a topological term, which, ignoring boundary terms, is of the form

$$\frac{11\, f_0}{60\,\pi^2}\int_M R^* R^* \sqrt{g}\, d^4x. \tag{1.757}$$

There is then an additional contribution from the fourth term of (1.753). Using (1.715), this gives

$$-\frac{f_0}{48\,\pi^2}\, R \operatorname{Tr}((D^{(0,1)})^2) = -\frac{f_0}{12\,\pi^2}\, a\, R\,|\varphi|^2 - \frac{f_0}{24\,\pi^2}\, c\, R\,. \tag{1.758}$$

For a physical explanation of the term $R\,|\varphi|^2$ see [**131**] equation 10.3.3.

Higgs minimal coupling. These terms are given by

$$\frac{f_0}{8\,\pi^2}\,\mathrm{Tr}(\mathbb{M}(D_\mu\,\varphi)^2) = \frac{f_0}{2\,\pi^2}\,a\,|D_\mu\,\varphi|^2, \tag{1.759}$$

where we used (1.743) and (1.753).

Higgs mass terms. There are two contributions with opposite signs. Using (1.715), one sees that the second term in (1.753) gives a term

$$-\frac{2\,f_2}{\pi^2}\,\Lambda^2\,a\,|\varphi|^2\,. \tag{1.760}$$

On the other hand, using (1.716), we obtain from the fourth term in (1.753) the additional term

$$\frac{f_0}{8\,\pi^2}\,8\,e\,|\varphi|^2 = \frac{e\,f_0}{\pi^2}\,|\varphi|^2. \tag{1.761}$$

Thus, combining (1.760) and (1.761), one obtains a Higgs mass term of the form

$$\frac{1}{\pi^2}(-2\,a\,f_2\,\Lambda^2\,+\,e\,f_0)\,|\varphi|^2\,. \tag{1.762}$$

Higgs quartic potential. The only contribution, in this case, comes from the fourth term in (1.753), i.e. from the term

$$\frac{f_0}{8\,\pi^2}\,\mathrm{Tr}((D^{0,1})^4).$$

Using (1.716), this gives the term

$$\frac{f_0}{2\,\pi^2}\,b\,|\varphi|^4\,. \tag{1.763}$$

Yang–Mills terms. For the Yang–Mills terms the computation is the same as in §11.4. In particular, by the same argument given in §11.4 we obtain the coefficient $\frac{f_0}{24\pi^2}$ in front of the trace of the square of the curvature.

The term $G^i_{\mu\nu}\,\overline{G}^{\mu\nu i}$, which corresponds to the gluons, has an additional coefficient $3\times 4\times 2 = 24$; since there are three generations, 4 quarks per generation (u_R, d_R, u_L, d_L), and a factor of two coming from the matter and anti-matter sectors $\mathcal{H}_f$ and $\mathcal{H}_{\bar{f}}$. Thus, because of the presence of the coefficient $\frac{g_3}{2}$, we get

$$\frac{f_0\,g_3^2}{4\pi^2}\,\mathrm{Tr}(G_{\mu\nu}\,\overline{G}^{\mu\nu}) = 2\,\frac{f_0\,g_3^2}{4\pi^2}\,G^i_{\mu\nu}\,\overline{G}^{\mu\nu}_i = \frac{f_0\,g_3^2}{2\pi^2}\,G^i_{\mu\nu}\,\overline{G}^{\mu\nu}_i\,,$$

where we use the property (1.422) for the Gell-Mann matrices.

For the weak interaction bosons W^α the argument is similar and one gets an additional coefficient $3\times 4\times 2 = 24$, here 3 is the number of generations, $4 = 3+1$ accounts for the three colors of quarks and one of leptons per isodoublet and per generation $(u_L = (u_j)_L, d_L = (d_j)_L, \nu_L, e_L)$, and the factor of 2 corresponds to the matter and anti-matter sectors $\mathcal{H}_f$ and $\mathcal{H}_{\bar{f}}$. Thus, using the property $\mathrm{Tr}(\sigma_a\sigma_b) = 2\delta_{ab}$ of the Pauli matrices, we obtain the similar term

$$\frac{f_0\,g_2^2}{4\pi^2}\,\mathrm{Tr}(F_{\mu\nu}\,\overline{F}^{\mu\nu}) = 2\,\frac{f_0\,g_2^2}{4\pi^2}\,F^a_{\mu\nu}\,\overline{F}^{\mu\nu}_a = \frac{f_0\,g_2^2}{2\pi^2}\,F^a_{\mu\nu}\,\overline{F}^{\mu\nu}_a\,.$$

Finally, for the hypercharge generator B_μ we obtain the additional coefficient

$$2\times\left(\left((\frac{4}{3})^2+(\frac{2}{3})^2+2(\frac{1}{3})^2\right)\times 3+(2^2+2)\right)\times 3 = 80, \tag{1.764}$$

which comes from the hypercharges as in Proposition 1.207, see (1.737) and (1.738). This gives an additional coefficient of $\frac{10}{3}$ in the corresponding term

$$\frac{10}{3}\frac{f_0\,g_1^2}{4\pi^2}\,B_{\mu\nu}\,\overline{B}^{\mu\nu} = \frac{5}{3}\,\frac{f_0\,g_1^2}{2\pi^2}\,B_{\mu\nu}\,\overline{B}^{\mu\nu}.$$

This completes the proof of Theorem 1.212. □

16.2. Fermionic action and Pfaffian.

The spectral action functional (1.526) applied to the operator D_A with inner fluctuations, as considered above, accounts for the bosonic part of the action. We now explain how one obtains the fermionic part.

The KO-dimension of the finite space F is $6 \in \mathbb{Z}/8$, hence the KO-dimension of the product geometry $M \times F$ (for M a spin 4-manifold) is now $2 \in \mathbb{Z}/8$. This means that, according to Definition 1.124, the commutation rules between D, J and γ are of the form

$$J^2 = -1, \quad JD = DJ, \quad \text{and} \quad J\gamma = -\gamma J\,. \tag{1.765}$$

We let $\mathcal{H}^+$ denote the *even part* of the Hilbert space $\mathcal{H}$, namely

$$\mathcal{H}^+ = \{\xi \in \mathcal{H} \mid \gamma\,\xi = \xi\}. \tag{1.766}$$

The commutation relations (1.765) yield a natural construction of an antisymmetric bilinear form on $\mathcal{H}^+$, in the following way.

PROPOSITION 1.213. *Let $(\mathcal{A}, \mathcal{H}, D, J, \gamma)$ be a real even spectral triple of KO-dimension $2 \in \mathbb{Z}/8$. Let $\mathcal{H}^+ \subset \mathcal{H}$ be the even part as in* (1.766)*. Then the following hold.*

(1) *The expression*

$$\mathfrak{A}_D(\xi', \xi) = \langle J\,\xi', D\,\xi\rangle\,, \quad \forall \xi, \xi' \in \mathcal{H}^+ \tag{1.767}$$

defines an antisymmetric bilinear form on $\mathcal{H}^+ = \{\xi \in \mathcal{H}\,,\ \gamma\,\xi = \xi\}$.

(2) *The trilinear pairing* (1.767) *between D, ξ and ξ' is gauge invariant under the adjoint action of the unitary group of $\mathcal{A}$, namely*

$$\mathfrak{A}_D(\xi', \xi) = \mathfrak{A}_{D_u}(\mathrm{Ad}(u)\xi', \mathrm{Ad}(u)\xi)\,, \qquad \textit{with} \quad D_u = \mathrm{Ad}(u)\,D\,\mathrm{Ad}(u^*). \tag{1.768}$$

PROOF. (1) We use an inner product which is antilinear in the first variable. Thus, since J is antilinear, $\mathfrak{A}$ is a bilinear form. Let us check that $\mathfrak{A}$ is antisymmetric. One has

$$\begin{aligned} \mathfrak{A}_D(\xi, \xi') = & \quad \langle J\,\xi, D\,\xi'\rangle &= -\langle J\,\xi, J^2\,D\,\xi'\rangle \\ = & \ -\langle J\,D\,\xi',\,\xi\rangle &= -\langle D\,J\,\xi',\,\xi\rangle \\ = & \ -\langle J\,\xi',\,D\,\xi\rangle, \end{aligned}$$

where we used the unitarity of J, i.e. the equality

$$\langle J\,\xi, J\,\eta\rangle = \langle \eta, \xi\rangle\,, \quad \forall \xi, \eta \in \mathcal{H}. \tag{1.769}$$

Finally, one can restrict the antisymmetric form $\mathfrak{A}_D$ to $\mathcal{H}^+$ without automatically getting zero, since one has

$$\gamma\,JD = \ JD\,\gamma\,.$$

(2) Let us check that $\mathrm{Ad}(u)$ commutes with J. By definition $\mathrm{Ad}(u) = u\,(u^*)^0 = u\,JuJ^{-1}$. Thus, we have

$$J\,\mathrm{Ad}(u) = J\,uJuJ^{-1} = u\,J\,uJJ^{-1} = uJu = \mathrm{Ad}(u)\,J\,,$$

where we used the commutation of u with JuJ. Since $\mathrm{Ad}(u)$ is unitary, one gets (1.768). ☐

Definition 1.214. *The Pfaffian of an antisymmetric bilinear form $\mathfrak{A}$ is expressed in terms of the functional integral involving anticommuting "classical fermions" (cf. [**250**], §5.1). At the formal level, this means that we write*

$$\mathrm{Pf}(\mathfrak{A}) = \int e^{-\frac{1}{2}\mathfrak{A}(\tilde{\xi})}\,\mathcal{D}[\tilde{\xi}]. \tag{1.770}$$

We define

$$\mathcal{H}^+_{cl} = \{\tilde{\xi}\,|\,\xi \in \mathcal{H}^+\} \tag{1.771}$$

to be the space of classical fermions (Graßmann variables) corresponding to $\mathcal{H}^+$ of (1.766).

Notice that $\mathfrak{A}(\xi,\xi) = 0$ when applied to vectors ξ, while $\mathfrak{A}(\tilde{\xi},\tilde{\xi}) \neq 0$ when applied to anticommuting Graßmann variables $\tilde{\xi}$.

Example 1.215. As the simplest example let us consider a two-dimensional vector space E with basis e_j and the antisymmetric bilinear form

$$\mathfrak{A}(\xi',\xi) = a(\xi'_1\xi_2 - \xi'_2\xi_1)$$

For $\tilde{\xi}_1$ anticommuting with $\tilde{\xi}_2$, using the basic rule

$$\int \tilde{\xi}_j\,d\tilde{\xi}_j = 1$$

(cf. [**250**], §5.1), one obtains

$$\int e^{-\frac{1}{2}\mathfrak{A}(\tilde{\xi})}\,\mathcal{D}[\tilde{\xi}] = \int e^{-a\,\tilde{\xi}_1\tilde{\xi}_2}\,d\tilde{\xi}_1 d\tilde{\xi}_2 = a.$$

In order to incorporate the fermionic part in the action, we consider, instead of the spectral action functional (1.526), the following modification.

Definition 1.216. *Suppose given a real even spectral triple $(\mathcal{A},\mathcal{H},D,J,\gamma)$ of KO-dimension 2 mod 8. For such a spectral triple the spectral action functional with fermions is the functional*

$$\mathrm{Tr}(f(D_A/\Lambda)) + \frac{1}{2}\langle\,J\,\tilde{\xi},D_A\,\tilde{\xi}\rangle. \tag{1.772}$$

where $\tilde{\xi} \in \mathcal{H}^+_{cl}$ and D_A is the Dirac operator with the unimodular inner fluctuations A.

Notice how the action functional (1.772) depends explicitly on all the data $(\mathcal{A},\mathcal{H},D,J,\gamma)$ of the spectral triple. It is the functional (1.772) that recovers the full Lagrangian of the Standard Model minimally coupled to gravity.

16.3. Fermion doubling, Pfaffian and Majorana fermions.

The use of the Pfaffian as a square root of the determinant solves the fermion doubling problem in the previous formulations of the Standard Model via noncommutative geometry, which was pointed out in [**210**]. The solution obtained by a better choice of the KO-dimension of the space F and hence of $M \times F$ is related to the point made in [**149**].

The point about the "fermion doubling problem" is that, if one introduces fermions in the spectral action using the pairing $\langle \bar{\psi}, D_A \psi \rangle$ instead of the pairing $\langle J\tilde{\psi}, D_A \tilde{\psi} \rangle$ that we introduced in the previous section, one obtains a number of fermion degrees of freedom that is equal to four times what it should be. The reason is that one includes one Dirac fermion for each of the chiral degrees of freedom L and R as well as the mirror fermion in $\mathcal{H}_{\bar{f}}$ for each fermion in $\mathcal{H}_f$.

The advantage of working with $\langle J\tilde{\psi}, D_A \tilde{\psi} \rangle$ is that, when considering the functional integral for the action functional (1.772), this has the effect of dividing by 4 the number of degrees of freedom by taking a fourth root of a determinant.

In fact, by Proposition 1.213 we are dealing with an antisymmetric bilinear form and the functional integral involving anticommuting Graßmann variables delivers the Pfaffian, which takes care of a square root. Again by Proposition 1.213, we can restrict the functional integration to the chiral subspace $\mathcal{H}^+_{cl}$ of (1.771), hence gaining another factor of two.

To be more explicit, we spell out in detail what happens. Consider the case of the quarks. With the basis $q_L, q_R, \bar{q}_L, \bar{q}_R$ in $\mathcal{H}_F$, in the reduction to $\mathcal{H}_+$ we write a generic vector as

$$\zeta = \xi_L \otimes q_L + \xi_R \otimes q_R + \eta_R \otimes \bar{q}_L + \eta_L \otimes \bar{q}_R, \tag{1.773}$$

where the subscripts L and R indicate the chirality of the usual spinors $\xi, \eta \in L^2(M, S)$. Similarly, one has

$$J\,\zeta' = J_M \xi'_L \otimes \bar{q}_L + J_M \xi'_R \otimes \bar{q}_R + J_M \eta'_R \otimes q_L + J_M \eta'_L \otimes q_R \tag{1.774}$$

and

$$\begin{aligned} \zeta'' = (\partial\!\!\!/_M \otimes 1)\, J\,\zeta' = \;& \partial\!\!\!/_M\, J_M \xi'_L \otimes \bar{q}_L + \partial\!\!\!/_M\, J_M \xi'_R \otimes \bar{q}_R \\ +\;& \partial\!\!\!/_M\, J_M \eta'_R \otimes q_L + \partial\!\!\!/_M\, J_M \eta'_L \otimes q_R. \end{aligned} \tag{1.775}$$

Thus, since the operator $\partial\!\!\!/_M\, J_M$ anticommutes with γ_5 in $L^2(M, S)$, we see that the vector ξ'' still belongs to $\mathcal{H}_+$, i.e. it is of the form (1.773). One gets then

$$\begin{aligned} \langle (\partial\!\!\!/_M \otimes 1)\, J\,\zeta', \zeta \rangle = \;& \langle \partial\!\!\!/_M\, J_M \xi'_L, \eta_R \rangle + \langle \partial\!\!\!/_M\, J_M \xi'_R, \eta_L \rangle \\ +\;& \langle \partial\!\!\!/_M\, J_M \eta'_R, \xi_L \rangle + \langle \partial\!\!\!/_M\, J_M \eta'_L, \xi_R \rangle. \end{aligned} \tag{1.776}$$

Upon writing spinors in the form $\xi = \xi_L + \xi_R$, the right-hand side can be written as

$$\langle (\partial\!\!\!/_M \otimes 1)\, J\,\zeta', \zeta \rangle = \langle \partial\!\!\!/_M\, J_M \xi', \eta \rangle + \langle \partial\!\!\!/_M\, J_M \eta', \xi \rangle. \tag{1.777}$$

This is an antisymmetric bilinear form in $L^2(M, S) \oplus L^2(M, S)$. Indeed if $\zeta' = \zeta$, i.e. $\xi' = \xi$ and $\eta' = \eta$ one has

$$\langle \partial\!\!\!/_M\, J_M \xi, \eta \rangle = -\langle \partial\!\!\!/_M\, J_M \eta, \xi \rangle, \tag{1.778}$$

since J_M commutes with $\partial\!\!\!/_M$ and has square -1.

Notice that it is because of (1.778) that one needs the factor $\frac{1}{2}$ in front of the fermionic action in (1.772), since in the Dirac sector the same expression repeats itself twice.

At the level of the fermionic functional integral the classical fermions $\tilde{\xi}$ and $\tilde{\eta}$ anticommute. Thus, up to the factor 2 taken care of by the $\frac{1}{2}$ in front of the fermionic term, one gets

$$\int e^{\langle J_M \tilde{\eta},\, \partial\!\!\!/_M\, \tilde{\xi}\rangle}\, \mathcal{D}[\tilde{\eta}]\mathcal{D}[\tilde{\xi}],$$

where $\tilde{\xi}$ and $\tilde{\eta}$ are independent anticommuting variables, with the same notation as in (1.771).

This coincides with the prescription for the Euclidean functional integral given in [**61**] (see "The use of instantons", §5.2) when using J_M to identify $L^2(M, S)$ with its dual.

The Dirac Yukawa terms simply replace $\partial\!\!\!/_M \otimes 1$ in the expression above by an operator of the form

$$\partial\!\!\!/_M \otimes 1 + \gamma_5 \otimes T,$$

where $T = T(x)$ acts as a matrix-valued function on the bundle $S \otimes \mathcal{H}_F$.

By construction, T preserves $\mathcal{H}_f$ and anticommutes with γ_F. Thus, one gets an equation of the form

$$\begin{aligned}(\gamma_5 \otimes T) J\,\zeta' = {} & T_1\, J_M \xi'_L \otimes \bar{q}_R + T_2\, J_M \xi'_R \otimes \bar{q}_L \\ + {} & T_3\, J_M \eta'_R \otimes q_R + T_4\, J_M \eta'_L \otimes q_L,\end{aligned}$$

where the T_j are endomorphisms of the spinor bundle commuting with the γ_5 matrix. In particular, the result is a vector in $\mathcal{H}^+$. Thus, one gets

$$\begin{aligned}\langle (\gamma_5 \otimes T)\, J\,\zeta', \zeta\rangle = {} & \langle T_1\, J_M \xi'_L, \eta_L\rangle + \langle T_2\, J_M \xi'_R, \eta_R\rangle \\ + {} & \langle T_3\, J_M \eta'_R, \xi_R\rangle + \langle T_4\, J_M \eta'_L, \xi_L\rangle.\end{aligned} \tag{1.779}$$

The expression (1.777) remains valid for the Dirac operator with Yukawa couplings, with the $J_M \xi'$ and $J_M \eta'$ on the left paired with the η and ξ, respectively. Thus, the Pfaffian of the corresponding classical fermions as Graßmann variables delivers the determinant of the Dirac operator.

We now come to the contribution of the piece of the operator D which, in the subspace $\nu_R, \bar{\nu}_R$, is of the form

$$T = \begin{pmatrix} 0 & \Upsilon_R^* \\ \Upsilon_R & 0 \end{pmatrix}$$

where Υ_R is a symmetric matrix in the flavor space, as in (1.669).

We use (1.773) and (1.774), replacing quarks by leptons, and we take a basis in which the matrix Υ_R is diagonal. We denote the corresponding eigenvalues by υ_R. We have

$$\zeta = \xi_L \otimes \nu_L + \xi_R \otimes \nu_R + \eta_R \otimes \bar{\nu}_L + \eta_L \otimes \bar{\nu}_R$$
$$J\,\zeta' = J_M \xi'_L \otimes \bar{\nu}_L + J_M \xi'_R \otimes \bar{\nu}_R + J_M \eta'_R \otimes \nu_L + J_M \eta'_L \otimes \nu_R$$

so that

$$(\gamma_5 \otimes T)\, J\,\zeta' = \gamma_5 \Upsilon_R J_M \xi'_R \otimes \nu_R + \gamma_5 \Upsilon_R J_M \eta'_L \otimes \bar{\nu}_R$$

$$\langle (\gamma_5 \otimes T)\, J\,\zeta', \zeta\rangle = \upsilon_R \langle \gamma_5 J_M \xi'_R, \xi_R\rangle + \upsilon_R \langle \gamma_5 J_M \eta'_R, \eta_R\rangle \tag{1.780}$$

The only effect of the γ_5 is an overall sign. The charge conjugation operator J_M is now playing a key role in the terms (1.780), where it defines an antisymmetric bilinear form on spinors of a given chirality (here right-handed ones).
Thus, in the Majorana sector we get a factor $\frac{1}{2}$ in front of the kinetic term. This corresponds to equation (4.20) of [**233**]. For the treatment of Majorana fermions in Euclidean functional integrals and the use of Pfaffians see e.g. [**171**], [**237**].
Notice that solving the fermion doubling problem using the pairing $\langle J\tilde{\psi}, D_A\tilde{\psi}\rangle$ in the functional (1.772) instead of $\langle \psi, D_A\psi\rangle$ is possible only because in this model the KO-dimension of the geometry $M \times F$ is 2 mod 8 instead of 4 mod 8 as in [**70**]. In Minkowskian signature, a similar solution of the fermion doubling problem was given in [**13**].

17. The Standard Model Lagrangian from the spectral action

In this section we show, following [**52**], how to recover the Standard Model Lagrangian of §9.4, with the modified terms of §9.6.2 that account for the mixing matrix and the Majorana mass terms for neutrinos, from the asymptotic expansion of the spectral action functional for the product geometry $M \times F$.

By covariance, if we want to show that the spectral action naturally recovers the Lagrangian for matter minimally coupled with gravity, i.e. with the Standard Model Lagrangian on a curved (Euclidean) spacetime, it is sufficient to check that we obtain the full Standard Model Lagrangian in flat spacetime, which is (up to switching back to Minkowskian signature) the Lagrangian considered in §9.4 and §9.6.2. Thus, for the rest of this section, in the explicit computation of the Lagrangian, we restrict to the flat case only.

The main theorem of this section is the following result.

THEOREM 1.217. *Let M be a compact Riemannian spin 4-manifold and F the finite noncommutative geometry of KO-dimension 6 described above. Let $M \times F$ be the product geometry.*

1. *The unimodular subgroup of the unitary group acting by the adjoint representation $\mathrm{Ad}(u)$ on $\mathcal{H}$ is the group of gauge transformations of SM.*
2. *The unimodular inner fluctuations of the metric give the gauge bosons of SM.*
3. *The action of the full Standard Model (with neutrino mixing and see-saw mechanism) minimally coupled to Einstein gravity in Euclidean form is given by the action functional (1.772) of Definition 1.216.*

PROOF. (1) and (2) follow from the results of §15.1 and §15.2 and 15.4 above.

(3) We subdivide the proof of the theorem into various subsections, where we obtain from the action functional (1.772) the various parts of the Standard Model Lagrangian as in (1.423).

- Yukawa coupling $\mathcal{L}_{Hf}$ as in (1.442)
- Gauge fermion couplings $\mathcal{L}_{gf}$ as in (1.443)
- Higgs self-coupling $\mathcal{L}_H$ as in (1.444)
- Self-coupling of gauge fields $\mathcal{L}_g$ as in (1.445)
- Minimal coupling of Higgs fields $\mathcal{L}_{Hg}$ as in (1.446)

We begin by a convenient reformulation of the asymptotic expansion of the spectral action on $M \times F$ obtained in Theorem 1.212.

17.1. Change of variables in the asymptotic formula and unification. We first perform a trivial rescaling of the Higgs field φ so that kinetic terms are normalized. To normalize the Higgs fields kinetic energy we have to rescale φ to

$$\mathbf{H} = \frac{\sqrt{a\, f_0}}{\pi}\varphi\,, \tag{1.781}$$

so that the kinetic term becomes

$$\int_M \frac{1}{2}|D_\mu \mathbf{H}|^2 \sqrt{g}\, d^4x. \tag{1.782}$$

We now introduce a relation between the parameters of the model. This corresponds to the relation between the coupling constants at unification, which is common to all the grand unified theories. The relation we impose here is dictated by the normalization of the kinetic terms that we discuss in detail in Lemma 1.228 below. Namely, we impose a relation between the coupling constants g_1, g_2, g_3 and the coefficient f_0, of the form

$$\frac{g_3^2\, f_0}{2\pi^2} = \frac{1}{4} \qquad \text{and} \qquad g_3^2 = g_2^2 = \frac{5}{3}\, g_1^2\,. \tag{1.783}$$

LEMMA 1.218. *Under the change of variable* (1.781) *and with the condition* (1.783), *the bosonic action* (1.754) *takes the form*

$$\begin{aligned} S = &\int_M \Big(\tfrac{1}{2\kappa_0^2} R + \alpha_0 C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma} + \gamma_0 + \tau_0 R^* R^* \\ + &\tfrac{1}{4} G^i_{\mu\nu} \overline{G}^{\mu\nu i} + \tfrac{1}{4} F^\alpha_{\mu\nu} \overline{F}^{\mu\nu\alpha} + \tfrac{1}{4} B_{\mu\nu} \overline{B}^{\mu\nu} \\ + &\tfrac{1}{2} |D_\mu \mathbf{H}|^2 - \mu_0^2 |\mathbf{H}|^2 - \xi_0 R |\mathbf{H}|^2 + \lambda_0 |\mathbf{H}|^4 \Big) \sqrt{g}\, d^4x, \end{aligned} \tag{1.784}$$

where

$$\begin{aligned} \tfrac{1}{\kappa_0^2} = &\ \tfrac{96\, f_2\, \Lambda^2 - f_0\, c}{12\, \pi^2} \\ \mu_0^2 = &\ 2\, \tfrac{f_2\, \Lambda^2}{f_0} - \tfrac{e}{a} \\ \alpha_0 = &\ -\tfrac{3\, f_0}{10\, \pi^2} \\ \tau_0 = &\ \tfrac{11\, f_0}{60\, \pi^2} \\ \gamma_0 = &\ \tfrac{1}{\pi^2} (48\, f_4\, \Lambda^4 - f_2\, \Lambda^2\, c + \tfrac{f_0}{4}\, d) \\ \lambda_0 = &\ \tfrac{\pi^2}{2\, f_0} \tfrac{b}{a^2} \\ \xi_0 = &\ \tfrac{1}{12} \end{aligned} \tag{1.785}$$

PROOF. This follows directly from Theorem 1.212, using (1.781) and (1.783). □

Imposing the condition (1.783) has the effect of replacing the term

$$\frac{f_0}{2\,\pi^2}(g_3^2 G^i_{\mu\nu} \overline{G}^{\mu\nu i} + g_2^2 F^\alpha_{\mu\nu} \overline{F}^{\mu\nu\alpha} + \frac{5}{3} g_1^2 B_{\mu\nu} \overline{B}^{\mu\nu})$$

of Theorem 1.212 with the normalized expression

$$\frac{1}{4} G^i_{\mu\nu} \overline{G}^{\mu\nu i} + \frac{1}{4} F^\alpha_{\mu\nu} \overline{F}^{\mu\nu\alpha} + \frac{1}{4} B_{\mu\nu} \overline{B}^{\mu\nu},$$

which will match the gauge bosons self-interaction terms of the Standard Model Lagrangian (cf. Lemma 1.228 below). The fact of imposing (1.783) means that the action functional (1.784) is considered as the *bare action* at *unification scale* Λ, i.e. where the merging $g_3^2 = g_2^2 = 5g_1^2/3$ of the coupling constants supposedly takes place. The relation $g_3^2 = g_2^2 = 5g_1^2/3$ coincides with that obtained in grand unification theories (cf. [**53**] and [**233**] §9). We recall briefly in §17.2 below how the unification scale is computed.

COROLLARY 1.219. *The change of notations* (1.781) *for the Higgs fields can be reformulated in the form*

$$\mathbf{H} = \frac{1}{\sqrt{2}}\frac{\sqrt{a}}{g}(1+\psi) = (\frac{2M}{g} + H - i\phi^0, -i\sqrt{2}\phi^+)\,. \tag{1.786}$$

PROOF. As in §9.4, we let g denote the weak coupling constant $g = g_2$, with $\alpha = s_w^2 g^2/(4\pi)$ the fine structure constant, and we let M denote the mass of the W. The result then follows using (1.783) to replace $\frac{\sqrt{a\,f_0}}{\pi}$ by $\frac{1}{\sqrt{2}}\frac{\sqrt{a}}{g}$. ☐

In particular, the change of variables (1.786) sets the the mass M of the W to be

$$2M = \sqrt{\frac{a}{2}}, \tag{1.787}$$

with a as in (1.717). The vacuum expectation value of the Higgs field is

$$v = \frac{2M}{g}. \tag{1.788}$$

17.2. Coupling constants at unification.

As in (1.420), we set $\alpha_i = g_i^2/(4\pi)$, with $\alpha_{\rm em} = s_w^2\alpha_2$ the fine structure constant, for $s_w = \sin(\theta_w)$. The infrared value is $\alpha_{\rm em} \sim 1/137.036$, but it is running as a function of the energy and increases to the value $\alpha_{\rm em}(M_Z) = 1/128.09$ already, at the energy $M_Z \sim 91.188$ GeV.

Assuming the "big desert" hypothesis, i.e. the absence of "new physics" (save for the Higgs) between current energy scales and unification, the running of the three couplings α_i is known. It is obtained via the renormalization group equation that we discussed in the first part of this chapter (see §6.6).

Considering only 1-loop corrections, the beta function (cf. Definition 1.46 and Corollary 1.49) is given by ([**187**], [**4**])

$$\beta_{g_i} = (4\pi)^{-2}\, b_i\, g_i^3, \qquad \text{with} \quad b = (\frac{41}{6}, -\frac{19}{6}, -7), \tag{1.789}$$

so that ([**255**]) the renormalization group equation is then of the form

$$\begin{aligned} \alpha_1^{-1}(\Lambda) &= \alpha_1^{-1}(M_Z) - \frac{41}{12\pi}\log\frac{\Lambda}{M_Z} \\ \alpha_2^{-1}(\Lambda) &= \alpha_2^{-1}(M_Z) + \frac{19}{12\pi}\log\frac{\Lambda}{M_Z} \\ \alpha_3^{-1}(\Lambda) &= \alpha_3^{-1}(M_Z) + \frac{42}{12\pi}\log\frac{\Lambda}{M_Z} \end{aligned} \tag{1.790}$$

where M_Z is the mass of the Z^0 vector boson.

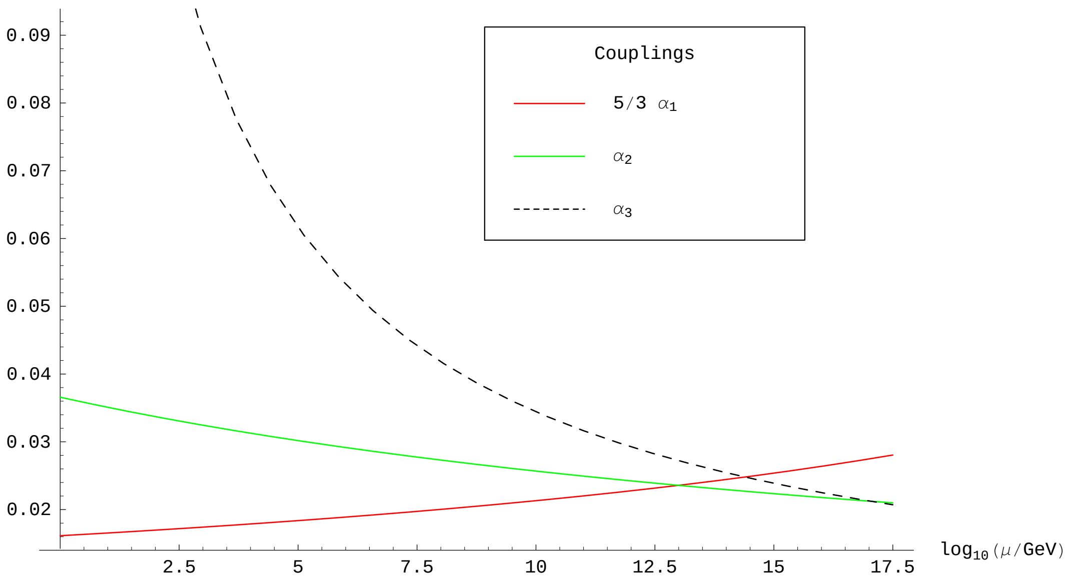


FIGURE 33. The running of the three couplings.

Notice that at 1-loop order the renormalization group equations for the coupling constants g_i are uncoupled from the other parameters of the Standard Model, but this is no longer the case when higher loop corrections are considered. For the explicit form of the beta function and renormalization group equation for 2-loop corrections see [**4**].

It is known that the predicted unification of the coupling constants does not hold exactly, which points to the existence of new physics, in contrast with the "big desert" hypothesis. Thus, one should bear in mind that the model discussed here is only approximate, in the same sense in which the prediction of all the grand unified theories that the coupling constants meet at the common value $g_3^2 = g_2^2 = 5g_1^2/3$ is known to be incompatible with the "big desert" hypothesis.
In fact, if one considers the actual experimental values, one has

$$g_1(M_Z) = 0.3575, \quad g_2(M_Z) = 0.6514, \quad g_3(M_Z) = 1.221, \tag{1.791}$$

one obtains the values

$$\alpha_1(M_Z) = 0.0101, \quad \alpha_2(M_Z) = 0.0337, \quad \alpha_3(M_Z) = 0.1186. \tag{1.792}$$

If one uses these values and runs the equation (1.790) one finds that the graphs of the running of the three constants α_i do not meet exactly, hence they do not specify a unique unification energy (cf. Figure 33 where the horizontal axis labels the logarithm in base 10 of the scale measured in GeV).

17.3. The coupling of fermions.

We now discuss how to obtain the parts $\mathcal{L}_{Hf}$ and $\mathcal{L}_{gf}$ as in (1.442) and (1.443) of the Standard Model Lagrangian. We first discuss the case of the minimal Standard Model with $\mathcal{L}_{Hf}^{\min}$ and $\mathcal{L}_{gf}^{\min}$ and then we show how to obtain the remaning terms for neutrino mixing and Majorana mass terms (see Lemma 1.221 below).

In the case of the minimal Standard Model, we have

$$
\begin{gathered}
\mathcal{L}_{Hf}^{\min} = \\
-\bar{e}^\lambda\, m_e^\lambda\, e^\lambda - \bar{u}_j^\lambda\, m_u^\lambda\, u_j^\lambda - \bar{d}_j^\lambda\, m_d^\lambda\, d_j^\lambda + \\
\tfrac{ig}{2\sqrt{2}}\tfrac{m_e^\lambda}{M}\left(-\phi^+(\bar{\nu}^\lambda(1-\gamma^5)e^\lambda) + \phi^-(\bar{e}^\lambda(1+\gamma^5)\nu^\lambda)\right) - \tfrac{g}{2}\tfrac{m_e^\lambda}{M}\left(H(\bar{e}^\lambda e^\lambda) + i\phi^0(\bar{e}^\lambda\gamma^5 e^\lambda)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^+\left(-m_d^\kappa(\bar{u}_j^\lambda C_{\lambda\kappa}(1-\gamma^5)d_j^\kappa) + m_u^\lambda(\bar{u}_j^\lambda C_{\lambda\kappa}(1+\gamma^5)d_j^\kappa)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^-\left(m_d^\lambda(\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger(1+\gamma^5)u_j^\kappa) - m_u^\kappa(\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger(1-\gamma^5)u_j^\kappa)\right) - \tfrac{g}{2}\tfrac{m_u^\lambda}{M}H(\bar{u}_j^\lambda u_j^\lambda) - \\
\tfrac{g}{2}\tfrac{m_d^\lambda}{M}H(\bar{d}_j^\lambda d_j^\lambda) + \tfrac{ig}{2}\tfrac{m_u^\lambda}{M}\phi^0(\bar{u}_j^\lambda\gamma^5 u_j^\lambda) - \tfrac{ig}{2}\tfrac{m_d^\lambda}{M}\phi^0(\bar{d}_j^\lambda\gamma^5 d_j^\lambda)
\end{gathered}
\tag{1.793}
$$

The matrix $C_{\lambda\kappa}$ is the CKM mixing matrix for quarks. It also enters the Lagrangian in the two gauge coupling terms where the down and up fermions are involved together in the expression

$$
\begin{gathered}
\mathcal{L}_{gf}^{\min} = \\
\tfrac{1}{2}ig_s\,\lambda_a^{ij}(\bar{q}_i^\sigma\gamma^\mu q_j^\sigma)g_\mu^a - \bar{e}^\lambda(\gamma\partial)e^\lambda - \bar{\nu}^\lambda\gamma\partial\nu^\lambda - \bar{u}_j^\lambda(\gamma\partial)u_j^\lambda - \bar{d}_j^\lambda(\gamma\partial)d_j^\lambda + \\
igs_wA_\mu\left(-(\bar{e}^\lambda\gamma^\mu e^\lambda) + \tfrac{2}{3}(\bar{u}_j^\lambda\gamma^\mu u_j^\lambda) - \tfrac{1}{3}(\bar{d}_j^\lambda\gamma^\mu d_j^\lambda)\right) + \tfrac{ig}{4c_w}Z_\mu^0\{(\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5)\nu^\lambda) + \\
(\bar{e}^\lambda\gamma^\mu(4s_w^2-1-\gamma^5)e^\lambda) + (\bar{d}_j^\lambda\gamma^\mu(\tfrac{4}{3}s_w^2-1-\gamma^5)d_j^\lambda) + (\bar{u}_j^\lambda\gamma^\mu(1-\tfrac{8}{3}s_w^2+\gamma^5)u_j^\lambda)\} + \\
\tfrac{ig}{2\sqrt{2}}W_\mu^+\left((\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5)e^\lambda) + (\bar{u}_j^\lambda\gamma^\mu(1+\gamma^5)C_{\lambda\kappa}d_j^\kappa)\right) + \\
\tfrac{ig}{2\sqrt{2}}W_\mu^-\left((\bar{e}^\lambda\gamma^\mu(1+\gamma^5)\nu^\lambda) + (\bar{d}_j^\kappa C_{\kappa\lambda}^\dagger\gamma^\mu(1+\gamma^5)u_j^\lambda)\right)
\end{gathered}
\tag{1.794}
$$

Since the matrix $C_{\lambda\kappa}$ is unitary the quadratic expressions in d_j^λ are unchanged by the change of variables given by

$$
d_{\lambda j} = C_{\lambda\kappa}\, d_j^\kappa\,, \qquad \bar{d}_{\lambda j} = \overline{C_{\lambda\kappa}}\,\bar{d}_j^\kappa = C_{\kappa\lambda}^\dagger\,\bar{d}_j^\kappa \tag{1.795}
$$

and in this way one can eliminate $C_{\lambda\kappa}$ in $\mathcal{L}_{gf}^{\min}$.

The term $\mathcal{L}_{gf}^{\min}$, which we now consider written in the new variables of (1.795), contains the kinetic terms of the fermions and their couplings to the various gauge fields. In the SU(3) case the coupling is simply given by the term

$$\frac{1}{2}ig_s\,\lambda_a^{ij}(\bar{q}_i^\sigma\gamma^\mu q_j^\sigma)g_\mu^a,$$

with λ the Gell-Mann matrices of (1.421) and $g_s = g_3$ the strong coupling. The coupling has a more complicated form in the case of the $(A, W^\pm, Z^0)$ fields. This involves, in particular, the hypercharges Y_L and Y_R assigned to the different fermions, quarks and leptons, and depending upon their chirality.

Recall here that the electromagnetic charge $Q_{\rm em}$ is related to the hypercharges Y_L and Y_R by the relations

$$2\,Q_{\rm em} = Y_R \quad \text{and} \quad 2\,Q_{\rm em} = Y_L + 2\,I_{3L}, \tag{1.796}$$

where I_{3L} is the third generator of the weak isospin group SU(2), as in (1.413). Thus, from the values of the hypercharges $Y = Y_L$ of §9.2 and §13.3, one sees that for $Q_{\rm em}$

one gets (for both right and left components) $\frac{2}{3}$ and $-\frac{1}{3}$ for u and d, respectively, and 0 and -1 for ν and e, respectively.
In order to relate the expressions $\mathcal{L}_{Hf}^{\min}$ and $\mathcal{L}_{gf}^{\min}$ of (1.793) and (1.794) to the spectral action functional (1.772), we first prove the following result.

LEMMA 1.220. *After the change of basis* (1.795) *on the down quarks, the following holds.*

(1) *The terms* $\mathcal{L}_{gf}^{\min}$ *are of the form*

$$\begin{aligned}\mathcal{L}_{gf}^{\min} &= -\sum_f [\overline{f}_L\gamma^\mu(\partial_\mu - ig\frac{\sigma_a}{2}W_{\mu a} - ig'\frac{Y_L}{2}B_\mu - ig''\frac{\lambda_b}{2}V_{\mu b})f_L \\ &+ \overline{f}_R\gamma^\mu(\partial_\mu - ig'\frac{Y_R}{2}B_\mu - ig''\lambda_b V_{\mu b})f_R]\end{aligned} \tag{1.797}$$

for both quarks and leptons, with $W_\mu^+ = \frac{W_{\mu 1} - iW_{\mu 2}}{\sqrt{2}}$, $W_\mu^- = \frac{W_{\mu 1} + iW_{\mu 2}}{\sqrt{2}}$ *and*

$$\begin{aligned} &g' = g\,\tan(\theta_w)\,, \qquad g'' = g_s\,, \\ &B_\mu = c_w\,A_\mu - s_w\,Z_\mu^0\,, \quad W_{\mu 3} = s_w\,A_\mu + c_w\,Z_\mu^0\end{aligned} \tag{1.798}$$

(2) *The terms* $\mathcal{L}_{Hf}^{\min}$ *are given, with the notation* (1.741), *by*

$$\mathcal{L}_{Hf}^{\min} = -\bar{f}\,T(0, K_e^{\min}, \varphi)\,f - \bar{f}\,T(K_u, K_d, \varphi)\,f \tag{1.799}$$

where

$$\varphi_1 = \frac{2M}{g} + H - i\phi^0, \qquad \varphi_2 = -i\sqrt{2}\phi^+ \tag{1.800}$$

and

$$\begin{aligned}(K_e^{\min})_{\mu\rho} &= \tfrac{g}{2M}m_e^\mu\delta_\mu^\rho \\ (K_u)_{\mu\rho} &= \tfrac{g}{2M}m_u^\mu\delta_\mu^\rho \\ (K_d)_{\mu\rho} &= \tfrac{g}{2M}m_d^\lambda C_{\mu\lambda}\delta_\lambda^\kappa C_{\kappa\rho}^\dagger.\end{aligned} \tag{1.801}$$

PROOF. (1) In Minkowski space a quark q is represented by a column vector and one has the relation

$$\bar{q} = q^*\,\gamma_0 \tag{1.802}$$

between q and $\bar{q}$. Thus, q and $\bar{q}$ have opposite chirality.
Since the γ^μ switch the chirality to its opposite and all the terms in (1.794) involve the γ^μ, they can be separated as a sum of terms only involving f_L and $\bar{f}_L$ and terms only involving f_R and $\bar{f}_R$. Since we are considering only the minimal Standard Model, the neutrinos ν^λ only appear as left-handed, i.e. as through the combination $(1+\gamma^5)\nu^\lambda$.
The last two lines of (1.794) correspond to the terms in $ig\frac{\sigma^a}{2}W_{\mu a}$ for the off-diagonal Pauli matrices $\sigma_1 = \sigma^a|_{a=1}$, $\sigma_2 = \sigma^a|_{a=2}$. The first line of (1.794) corresponds to the gluons and the kinetic terms. The terms involving the gluons g_μ^a in (1.794) give the strong coupling constant $g'' = g_s = g_3$. The second and third lines of (1.794) use the electromagnetic field A_μ, which is related to the U(1) gauge field B_μ by

$$g\,s_w\,(A_\mu - \tan(\theta_w)\,Z_\mu^0) = g'\,B_\mu. \tag{1.803}$$

This gives (notice the sign $- \times - = +$ in (1.797)) the terms (cf. (1.796))

$$ig'\frac{Y_R}{2}B_\mu = igs_w A_\mu Q_{\rm em} - ig\frac{s_w^2}{c_w}Z^0_\mu Q_{\rm em} \tag{1.804}$$

for the right-handed part, while on the left-handed sector one has

$$Q_{\rm em} = \frac{Y_L}{2} + \frac{\sigma_3}{2}$$

where the third Pauli matrix $\sigma_3 := \sigma^a|_{a=3}$ corresponds to the generator I_{3L} of the isospin SU(2), as in (1.796). The diagonal terms

$$ig\frac{\sigma_3}{2}W_{\mu 3} + ig'\frac{Y_L}{2}B_\mu$$

for the left-handed part are then of the form

$$ig\tfrac{\sigma_3}{2}W_{\mu 3} + igs_w(A_\mu - \tan(\theta_w)Z^0_\mu)(Q_{\rm em} - \tfrac{\sigma_1}{2}) =$$

$$igs_w A_\mu Q_{\rm em} - ig\tfrac{s_w^2}{c_w}Z^0_\mu Q_{\rm em} + (igW_{\mu 3} - igs_w(A_\mu - \tan(\theta_w)Z^0_\mu))\tfrac{\sigma_3}{2}.$$

The relation

$$(igW_{\mu 3} - ig\, s_w\,(A_\mu - \tan(\theta_w)\,Z^0_\mu)) = \frac{ig}{c_w}\,Z^0_\mu \tag{1.805}$$

then determines $W_{\mu 3}$ as a function of A_μ and Z^0_μ, in the form

$$W_{\mu 3} = \; s_w\,(A_\mu - \tan(\theta_w)\,Z^0_\mu) + \frac{1}{c_w}\,Z^0_\mu$$

which is more conveniently written as

$$W_{\mu 3} = s_w\,A_\mu + c_w\,Z^0_\mu. \tag{1.806}$$

The diagonal terms for the left-handed sector can then be written in the form

$$ig\,s_w\,A_\mu\,Q_{\rm em} - \;ig\,\frac{s_w^2}{c_w}\,Z^0_\mu\,Q_{\rm em} + \frac{ig}{c_w}\,Z^0_\mu\,\frac{\sigma_3}{2}. \tag{1.807}$$

This agrees with the factor $\frac{ig}{4c_w}$ multiplying $(1+\gamma^5)$ in (1.794). The term $(1+\gamma^5)$ is twice the projection on the left-handed particles. This accounts for one factor of 2, while the other one comes from the denominator in $\frac{\sigma_3}{2}$.
The term

$$\frac{ig}{4c_w}Z^0_\mu\{(\bar\nu^\lambda\gamma^\mu(1+\gamma^5)\nu^\lambda) + (\bar e^\lambda\gamma^\mu(4s_w^2 - 1 - \gamma^5)e^\lambda)\}$$

is fine. In fact, the neutrino has no electromagnetic charge, so that the terms with $Q_{\rm em}$ in (1.807) do not contribute, while one gets the term $-ig\,\frac{s_w^2}{c_w}\,Z^0_\mu\,Q_{\rm em}$ for the electron. The different signs of the $(1+\gamma^5)$ term come from the fact that the left-handed neutrino corresponds to the $+1$ component of σ_3 and the left-handed electron to the -1 component of σ_3. The other two terms

$$\frac{ig}{4c_w}Z^0_\mu\{(\bar d^\lambda_j\gamma^\mu(\frac{4}{3}s_w^2 - 1 - \gamma^5)d^\lambda_j) + (\bar u^\lambda_j\gamma^\mu(-\frac{8}{3}s_w^2 + 1 + \gamma^5)u^\lambda_j)\}$$

also match (1.807) correctly. In fact, the electromagnetic charge of the down quark is $-\frac{1}{3}$ and it has $\sigma_3 = -1$, while for the up quark the electromagnetic charge is $\frac{2}{3}$ and $\sigma_3 = 1$.

(2) Recall that by (1.802) q and $\bar{q}$ have opposite chiralities. Thus, when we spell out the various terms in terms of the chiral components, we always get combinations of the form $\bar{q}_L\, X\, q_R$ or $\bar{q}_R\, X\, q_L$.
We first look at the lepton sector. This gives

$$-\bar{e}^\lambda\, m_e^\lambda\, e^\lambda \quad +\tfrac{ig}{2\sqrt{2}}\tfrac{m_e^\lambda}{M}\left(-\phi^+(\bar{\nu}^\lambda(1-\gamma^5)e^\lambda)+\phi^-(\bar{e}^\lambda(1+\gamma^5)\nu^\lambda)\right)$$
$$-\tfrac{g}{2}\tfrac{m_e^\lambda}{M}\left(H(\bar{e}^\lambda e^\lambda)+i\phi^0(\bar{e}^\lambda\gamma^5 e^\lambda)\right).$$

The terms in $\bar{e}\, X\, e$ are of two types. The first gives

$$-\bar{e}^\lambda\, m_e^\lambda\,(1+\frac{g\,H}{2\,M})\, e^\lambda = -\bar{e}_L^\lambda\, m_e^\lambda\,(1+\frac{g\,H}{2\,M})\, e_R^\lambda - \bar{e}_R^\lambda\, m_e^\lambda\,(1+\frac{g\,H}{2\,M})\, e_L^\lambda.$$

The second type gives

$$-\frac{g}{2}\frac{m_e^\lambda}{M} i\phi^0(\bar{e}^\lambda\gamma^5 e^\lambda) = \frac{g}{2}\frac{m_e^\lambda}{M} i\phi^0(\bar{e}_L^\lambda e_R^\lambda) - \frac{g}{2}\frac{m_e^\lambda}{M} i\phi^0(\bar{e}_R^\lambda e_L^\lambda).$$

Thus, they combine together using the complex field

$$\psi_1 = H - i\phi^0 \tag{1.808}$$

to give

$$-\bar{e}_L^\lambda\, m_e^\lambda\,(1+\frac{g\,\psi_1}{2\,M})\, e_R^\lambda - \bar{e}_R^\lambda\, m_e^\lambda\,(1+\frac{g\,\bar{\psi}_1}{2\,M})\, e_L^\lambda.$$

Since we are dealing only with the minimal Standard Model, the terms where both e and ν appear involve only ν_L, hence only e_R. The fields $\phi^\pm$ are complex fields that are complex conjugates of each other. We let

$$\psi_2 = \,-\,i\,\sqrt{2}\,\phi^+. \tag{1.809}$$

The contribution of the terms involving both e and ν is then of the form

$$\bar{\nu}_L^\lambda\, m_e^\lambda\,(\frac{g\,\psi_2}{2\,M})\, e_R^\lambda + \bar{e}_R^\lambda\, m_e^\lambda\,(\frac{g\,\bar{\psi}_2}{2\,M})\,\nu_L^\lambda.$$

Notice that the field $\psi=(\psi_1,\psi_2)$, with ψ_1 and ψ_2 as in (1.808) and (1.809) differs from the ψ of (1.786) by a factor $2M/g$.

Let $T(M_1,M_2,\varphi)$ be defined as in (1.741). Combining the above, we see that we can express the terms of the lepton sector in $\mathcal{L}_{Hf}^{\min}$ in the form

$$-\,\bar{f}\,T(0,K_e^{\min},\varphi)\,f\,,\quad \varphi_1=\psi_1+\frac{2\,M}{g}\,,\quad \varphi_2=\psi_2, \tag{1.810}$$

where $K_e^{\min}$ is the diagonal matrix with diagonal entries the $\frac{g}{2M}\, m_e^\lambda$, as in (1.801).
We now consider the quark sector of $\mathcal{L}_{Hf}^{\min}$. This consists of the terms

$$-\bar{u}_j^\lambda\, m_u^\lambda\, u_j^\lambda - \bar{d}_j^\lambda\, m_d^\lambda\, d_j^\lambda +$$
$$\tfrac{ig}{2M\sqrt{2}}\phi^+\left(-m_d^\kappa(\bar{u}_j^\lambda C_{\lambda\kappa}(1-\gamma^5)d_j^\kappa)+m_u^\lambda(\bar{u}_j^\lambda C_{\lambda\kappa}(1+\gamma^5)d_j^\kappa)\right)+$$
$$\tfrac{ig}{2M\sqrt{2}}\phi^-\left(m_d^\lambda(\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger(1+\gamma^5)u_j^\kappa)-m_u^\kappa(\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger(1-\gamma^5)u_j^\kappa)\right)-\tfrac{g}{2}\tfrac{m_u^\lambda}{M}H(\bar{u}_j^\lambda u_j^\lambda)-$$
$$\tfrac{g}{2}\tfrac{m_d^\lambda}{M}H(\bar{d}_j^\lambda d_j^\lambda)+\tfrac{ig}{2}\tfrac{m_u^\lambda}{M}\phi^0(\bar{u}_j^\lambda\gamma^5 u_j^\lambda)-\tfrac{ig}{2}\tfrac{m_d^\lambda}{M}\phi^0(\bar{d}_j^\lambda\gamma^5 d_j^\lambda).$$

We write it in terms of the $d_{\lambda j}$ given by (1.795) instead of the d_j^λ. The terms of the form $\bar u\, X\, u$ are

$$-\bar u_j^\lambda\, m_u^\lambda\, u_j^\lambda - \frac{g}{2}\frac{m_u^\lambda}{M} H(\bar u_j^\lambda u_j^\lambda) + \frac{ig}{2}\frac{m_u^\lambda}{M}\phi^0(\bar u_j^\lambda \gamma^5 u_j^\lambda).$$

They are similar to the terms in $\bar e\, X\, e$ for the leptons but with an opposite sign in front of ϕ^0. Thus, if we let K_u be the diagonal matrix with diagonal entries the $\frac{g}{2M}\, m_u^\lambda$, we get the terms depending on φ_1 and K_u in the expression

$$-\,\bar f\; T(K_u, K_d, \varphi)\; f, \tag{1.811}$$

where K_d remains to be determined.
There are two other terms involving the m_u^λ, which are directly written in terms of the $d_{\lambda j}$. They are of the form

$$\frac{ig}{2M\sqrt{2}}\,\phi^+\, m_u^\lambda(\bar u_j^\lambda(1+\gamma^5)d_{\lambda j}) - \frac{ig}{2M\sqrt{2}}\,\phi^-\, m_u^\kappa(\bar{d_{\kappa j}}(1-\gamma^5)u_j^\kappa).$$

This is the same as

$$-\bar d_{\lambda jL}\, m_u^\lambda\, (\frac{g\,\bar\psi_2}{2\,M})\, u_{jR}^\lambda - \,\bar u_{jR}^\lambda\, m_u^\lambda\, (\frac{g\,\psi_2}{2\,M})\, d_{\lambda jL},$$

which corresponds to the other terms involving K_u in (1.811).
The remaining terms are

$$-\bar d_j^\lambda\, m_d^\lambda\, d_j^\lambda + \frac{ig\, m_d^\kappa}{2M\sqrt{2}}(-\phi^+\,(\bar u_j^\lambda C_{\lambda\kappa}(1-\gamma^5)d_j^\kappa) + \phi^-\,(\bar d_j^\lambda C^\dagger_{\lambda\kappa}(1+\gamma^5)u_j^\kappa)) \tag{1.812}$$
$$-\frac{g}{2}\frac{m_d^\lambda}{M}(H(\bar d_j^\lambda d_j^\lambda) + \,i\,\phi^0(\bar d_j^\lambda\gamma^5 d_j^\lambda)).$$

Except for the transition to the the $d_{\lambda j}$, these terms are the same as for the lepton sector. Thus, we define the matrix K_d in such a way that it satisfies

$$\bar d_{\lambda jL}\, K_d^{\lambda\kappa}\; d_{\kappa jR} + \bar d_{\lambda jR}\, K_d^{\dagger\lambda\kappa}\; d_{\kappa jL} = \;\frac{g}{2M}\, \bar d_j^\lambda\, m_d^\lambda\, d_j^\lambda.$$

We can just take the positive matrix obtained as the conjugate

$$(K_d)_{\mu\rho} = \frac{g}{2M} m_d^\lambda C_{\mu\lambda}\delta_\lambda^\kappa C^\dagger_{\kappa\rho} \tag{1.813}$$

as in (1.801).
The only terms that remain to be understood are then the cross terms (with up and down quarks) in (1.812). It might seem at first that one recognizes the expression for $d_{\lambda j} = C_{\lambda\kappa}\, d_j^\kappa$, but this is not the case. In fact, the summation index κ also appears elsewhere, namely in m_d^κ. One has

$$\frac{g}{2M} m_d^\kappa C_{\lambda\kappa} d_j^\kappa = \frac{g}{2M} m_d^\mu C_{\lambda\mu}\delta_\mu^\kappa C^\dagger_{\kappa\rho} d_{\rho j} = (K_d\, d)_{\lambda j}.$$

Thus, the cross terms in (1.812) can be written in the form

$$\frac{i}{\sqrt{2}}(-\phi^+\,(\bar u_j^\lambda K_d^{\lambda\kappa}(1-\gamma^5)d_{\kappa j}) + \phi^-\,(\bar d_{\lambda j}(K_d^\dagger)^{\lambda\kappa}(1+\gamma^5)u_j^\kappa)).$$

Thus, we get the complete expression (1.811). ∎

Now we consider the case of the Lagrangian $\mathcal{L}_{SM}$ with the additional terms for neutrino mixing and Majorana mass terms. The following result shows how to modify Lemma 1.220 to account for these additional terms.

LEMMA 1.221. *The neutrino masses and mixing are obtained in two additional steps. The first is the replacement*

$$T(0, K_e^{\min}, \varphi) \mapsto T(K_\nu, K_e, \varphi),$$

where the $K_e^{\min}$ *of* (1.801) *is replaced by*

$$(K_e)_{\lambda\kappa} = \frac{g}{2M} m_e^\mu C^{lep}_{\lambda\mu} \delta^\rho_\mu (C^{lep})^\dagger_{\rho\kappa}, \tag{1.814}$$

while K_ν *is the neutrino Dirac mass matrix*

$$(K_\nu)_{\lambda\kappa} = \frac{g}{2M} m_\nu^\lambda \delta^\kappa_\lambda. \tag{1.815}$$

The second step is the addition of the Majorana mass term

$$\mathcal{L}_{mass} = -\frac{1}{4}\,\bar\nu_\lambda\, M^R_{\lambda\kappa}\,(1-\gamma^5)\hat\nu_\kappa \,-\frac{1}{4}\,\bar{\hat\nu}_\lambda\, \bar M^R_{\lambda\kappa}\,(1+\gamma^5)\nu_\kappa. \tag{1.816}$$

PROOF. After performing the inverse of the change of variables (1.795) for the leptons, using the matrix C^{lep} instead of the CKM matrix, the new Dirac Yukawa coupling terms for the leptons imply the replacement of

$$-\frac{g}{2}\frac{m_e^\lambda}{M}\left(H(\bar e^\lambda e^\lambda) + i\phi^0(\bar e^\lambda \gamma^5 e^\lambda)\right)$$

by the terms

$$-\frac{g}{2}\frac{m_\nu^\lambda}{M}H(\bar\nu^\lambda\nu^\lambda) - \frac{g}{2}\frac{m_e^\lambda}{M}H(\bar e^\lambda e^\lambda) + \frac{ig}{2}\frac{m_\nu^\lambda}{M}\phi^0(\bar\nu^\lambda\gamma^5\nu^\lambda) - \frac{ig}{2}\frac{m_e^\lambda}{M}\phi^0(\bar e^\lambda\gamma^5 e^\lambda)$$

and of the terms

$$\frac{ig}{2\sqrt2}\frac{m_e^\lambda}{M}\left(-\phi^+(\bar\nu^\lambda(1-\gamma^5)e^\lambda) + \phi^-(\bar e^\lambda(1+\gamma^5)\nu^\lambda)\right)$$

by

$$\begin{aligned}&\tfrac{ig}{2M\sqrt2}\phi^+\left(-m_e^\kappa(\bar\nu^\lambda C^{lep}_{\lambda\kappa}(1-\gamma^5)e^\kappa) + m_\nu^\lambda(\bar\nu^\lambda C^{lep}_{\lambda\kappa}(1+\gamma^5)e^\kappa)\right) + \\ &\tfrac{ig}{2M\sqrt2}\phi^-\left(m_e^\lambda(\bar e^\lambda (C^{lep})^\dagger_{\lambda\kappa}(1+\gamma^5)\nu^\kappa) - m_\nu^\kappa(\bar e^\lambda (C^{lep})^\dagger_{\lambda\kappa}(1-\gamma^5)\nu^\kappa)\right),\end{aligned}$$

The structure we obtain in the lepton sector is now identical to that of the quark sector. The result then follows from Lemma 1.220.

The Majorana mass terms are of the form (1.816), where the coefficient $\frac14$ instead of $\frac12$ comes from the fact that $(1-\gamma^5)$ is twice the chiral projection on the R-chirality. The mass matrix M^R is a symmetric 3×3 matrix in the flavor space. ☐

In order to compare the expressions for $\mathcal{L}_{Hf}$ and $\mathcal{L}_{gf}$ obtained from Lemma 1.220 and Lemma 1.221 with the fermionic part of the functional (1.772) of Definition 1.216, we still need to discuss the effect of passing from the Minkowskian to the Euclidean formulation.

The Lagrangian $\mathcal{L}_{SM}$ is written in Lorentzian signature. In order to understand the Euclidean version of the action considered above, we start by treating the simpler case of the free Dirac field.

The action for a free Dirac field is given in Minkowski space by the action functional associated to the Lagrangian

$$-\bar u\,\gamma\partial u\;-\bar u\, m\, u. \tag{1.817}$$

In Euclidean space the action functional becomes instead of the form (cf. [**61**], "The use of instantons", §5.2)

$$S = -\int \bar{\psi}\,(i\,\gamma^{\mu}\,\partial_{\mu} - im)\,\psi\,d^4x, \tag{1.818}$$

where the symbols ψ and $\bar{\psi}$ now stand for *classical* fermions, i.e. independent anti-commuting Graßmann variables.
Notice that, in (1.818), the gamma matrices γ^{μ} are self-adjoint and the presence of $i = \sqrt{-1}$ in the mass term is crucial to ensure that the Euclidean propagator is of the form

$$\frac{\not{p} + im}{p^2 + m^2}.$$

Back to our case, consider the Dirac operator D_A that incorporates the inner fluctuations. Recall that D_A is given by the sum of two terms

$$D_A = D^{(1,0)} + \gamma_5 \otimes D^{(0,1)}, \tag{1.819}$$

where $D^{(0,1)}$ is given by (1.714) and $D^{(1,0)}$ is of the form

$$D^{(1,0)} = \sqrt{-1}\,\gamma^{\mu}(\nabla^s_{\mu} + \mathbb{A}_{\mu}), \tag{1.820}$$

where ∇^s is the spin connection (cf. (1.580)), while the $\mathbb{A}_{\mu}$ are as in Proposition 1.207.

We have the following result, which relates the bilinear form $\langle J\xi', (D^{(1,0)} + i \otimes D^{(0,1)})\xi\rangle$ for the operator $D^{(1,0)} + i \otimes D^{(0,1)}$, which corresponds as in (1.818) and (1.817) above to what we obtained in Lemma 1.220, and the bilinear form $\langle J\psi', D_A\psi\rangle$ for the Dirac operator D_A of (1.772).

LEMMA 1.222. *The unitary operator*

$$U = e^{i\frac{\pi}{4}\gamma_5} \otimes 1$$

commutes with $\mathcal{A}$ and γ. One has $JU = U^*J$ and

$$U\,D_A\,U = D^{(1,0)} + i \otimes D^{(0,1)}. \tag{1.821}$$

PROOF. Since γ_5 anticommutes with the γ^{μ}, one has $D^{(1,0)}\,e^{i\frac{\pi}{4}\gamma_5} = e^{-i\frac{\pi}{4}\gamma_5}\,D^{(1,0)}$. Moreover

$$U\,(\gamma_5 \otimes D^{(0,1)})\,U = (\gamma_5\,e^{i\frac{\pi}{2}\gamma_5}) \otimes D^{(0,1)} = i\,\otimes D^{(0,1)}$$

∎

The result of Lemma 1.222 can be restated as the equality of antisymmetric bilinear forms

$$\langle JU\xi', D_A U\xi\rangle = \langle J\xi', (D^{(1,0)} + i \otimes D^{(0,1)})\xi\rangle. \tag{1.822}$$

17.4. The mass relation at unification.

Notice that in (1.784) the matrices Υ_u, Υ_d, Υ_ν and Υ_e are only relevant up to an overall scale. In fact, they only enter in the coupling of the Higgs with fermions and, because of the rescaling (1.781), only through the terms

$$k_x := \frac{\pi}{\sqrt{a\, f_0}}\, \Upsilon_x, \tag{1.823}$$

for $x \in \{\nu, e, u, d\}$. The k_x are dimensionless matrices by construction. In fact, the coefficient

$$a = \mathrm{Tr}(\Upsilon_\nu^* \Upsilon_\nu + \Upsilon_e^* \Upsilon_e + 3(\Upsilon_u^* \Upsilon_u + \Upsilon_d^* \Upsilon_d)) \tag{1.824}$$

of (1.717) has the physical dimension of a (mass)^2.

We then conclude the discussion of the $\mathcal{L}_{Hf}$ and $\mathcal{L}_{gf}$ terms by the following result.

PROPOSITION 1.223. *The terms $\mathcal{L}_{Hf}$ and $\mathcal{L}_{gf}$ of the Standard Model Lagrangian are obtained from the functional* (1.772) *after identifying the $(K_\nu)_{\mu\rho}$, $(K_e)_{\mu\rho}$, $(K_u)_{\mu\rho}$ and $(K_d)_{\mu\rho}$ of* (1.815) *and* (1.801), *respectively, with the k_ν, k_e, k_u, k_d of* (1.823) *and the mass matrix M^R of* (1.816) *with the Υ_R.*

PROOF. This follows directly by combining the results of Lemma 1.220, Lemma 1.221, and Lemma 1.222 above. □

There is an interesting direct consequence of the identification of the k_x of (1.823) with the K_x of (1.815) and (1.801), which provides a physical prediction of the model.

LEMMA 1.224. *The mass matrices of* (1.826) *satisfy the constraint*

$$\sum_\sigma (m_\nu^\sigma)^2 + (m_e^\sigma)^2 + 3\,(m_u^\sigma)^2 + 3\,(m_d^\sigma)^2 = \, 8\, M^2, \tag{1.825}$$

where the sum is over generations.

PROOF. It might seem at first sight that one can simply use the identification

$$\begin{aligned} (k_u)_{\sigma\kappa} &= \frac{g}{2M}\, m_u^\sigma\, \delta_\sigma^\kappa \\ (k_d)_{\sigma\kappa} &= \frac{g}{2M}\, m_d^\mu\, C_{\sigma\mu} \delta_\mu^\rho C_{\rho\kappa}^\dagger \\ (k_\nu)_{\sigma\kappa} &= \frac{g}{2M}\, m_\nu^\sigma\, \delta_\sigma^\kappa \\ (k_e)_{\sigma\kappa} &= \frac{g}{2M}\, m_e^\mu\, C_{\sigma\mu}^{lep} \delta_\mu^\rho (C^{lep})_{\rho\kappa}^\dagger \end{aligned} \tag{1.826}$$

imposed by Proposition 1.223 as a way to define the matrices k_x but this overlooks the fact that (1.823) and the form (1.824) of the coefficient a imply the constraint

$$\mathrm{Tr}(k_\nu^* k_\nu + k_e^* k_e + 3(k_u^* k_u + k_d^* k_d)) = \, 2\, g^2\,, \tag{1.827}$$

where we use (1.783) to replace $\frac{\pi^2}{f_0}$ by $2\, g^2$. Using (1.826), we then obtain the constraint (1.825), where summation is performed with respect to the flavor index σ. Notice that g^2 appears in the same way on both sides and drops out of the equation. □

We discuss the physical interpretation of the imposition of this constraint at unification scale.

The relation (1.825) means that, at unification scale where our model is considered, there is a simple relation between the masses of quarks and leptons and the mass of the W boson, which is expressed as a quadratic polynomial with integer coefficients.

We need to check that, when running down to ordinary energies the values of the masses through the renormalization group equation, the prediction made by (1.825) is compatible with the known physics of the Standard Model.

In order to do that, we first need to discuss another aspect of our model, which is related to the presence of the Υ_R term in the Dirac operator of the finite geometry, which accounts for the Majorana mass terms of $\mathcal{L}_{SM}$.

17.5. The see-saw mechanism.

In models of neutrino physics where neutrinos are described by Majorana fermions, the smallness of neutrino masses is explained through what is called the "see-saw mechanism". In fact, as we explained briefly in §9.6.1 above, the relative size of the terms in the matrix $\mathcal{S}_{\kappa\lambda}$ of (1.439) for Majorana fermions can be used to make one particle mass very small at the expense of having another one with very large mass (see e.g. §7.2.4 and §8.2.1 of [**233**]).

In our model, as we saw in Lemma 1.221 and Proposition 1.223 above, the Majorana mass terms are accounted for by the presence of the Υ_R term in the Dirac operator of the finite noncommutative geometry F. We now show how this provides the desired see-saw mechanism.

LEMMA 1.225. *Let $D = D(\Upsilon)$ be as in* (1.663). *The restriction of $D(\Upsilon)$ to the subspace of $\mathcal{H}_F$ with basis the $(\nu_R, \nu_L, \bar{\nu}_R, \bar{\nu}_L)$ is given by a matrix of the form*

$$\begin{pmatrix} 0 & M_\nu^* & M_R^* & 0 \\ M_\nu & 0 & 0 & 0 \\ M_R & 0 & 0 & \bar{M}_\nu^* \\ 0 & 0 & \bar{M}_\nu & 0 \end{pmatrix} \tag{1.828}$$

where $M_\nu = \frac{2M}{g} K_\nu$ with K_ν as in (1.815) *and M_R is the matrix $M^R_{\kappa\lambda}$ of the Majorana mass term* (1.816).

PROOF. This is simply rephrasing the explicit form of the Dirac operator $D(\Upsilon)$ of §13.4. In fact, we know by Proposition 1.223 that the matrix $M^R_{\kappa\lambda}$ is identified with the symmetric matrix Υ_R in $D(\Upsilon)$, while the term $M_\nu = \frac{2M}{g} K_\nu$ is nothing but the term Υ_ν, by (1.815) together with the fact that, by Proposition 1.223, K_ν is identified with the k_ν of (1.823), and we have the relation $2M = \sqrt{a/2}$ of (1.787), with a as in (1.824) from the change of variables (1.786) in the Higgs field. $\square$

The largest eigenvalue of M_R is set to the order of the unification scale by the equations of motion of the spectral action as in the following result.

LEMMA 1.226. *Assume that the matrix M_R is a multiple of a fixed matrix k_R, i.e. is of the form $M_R = x\, k_R$. In flat space, and assuming that the Higgs vacuum expectation value is negligible with respect to unification scale, the equations of motion* *of the spectral action fix x to be either $x = 0$ (unstable) or satisfying*

$$x^2 = \frac{2\, f_2\, \Lambda^2\, \mathrm{Tr}(k_R^* k_R)}{f_0\, \mathrm{Tr}((k_R^* k_R)^2)}. \tag{1.829}$$

PROOF. The value of x is fixed by the equations of motion of the spectral action

$$\partial_u \mathrm{Tr}(f(D_A/\Lambda)) = 0, \tag{1.830}$$

with $u = x^2$.

One can see from (1.717) that x only appears in the coefficients c, d, and e. In the variation (1.830), the terms in the spectral action (1.754) of Theorem 1.212 containing the coefficient c and e produce linear terms in x^2, proportional to the scalar curvature R and the square $|\varphi|^2$ of the Higgs vacuum expectation value, and an additional linear term coming from the cosmological term. The cosmological term also contains the coefficient d, which depends quadratically on x^2. In flat space, and under the assumption that $|\varphi|^2$ is sufficiently small, (1.830) then corresponds to minimizing the cosmological term.

This gives

$$\partial_x(-f_2\, \Lambda^2\, c + \frac{f_0}{4}\, d) = 0, \qquad c = x^2\, \mathrm{Tr}(k_R^* k_R), \qquad d = x^4\, \mathrm{Tr}((k_R^* k_R)^2). \tag{1.831}$$

Thus, we get $M_R = x k_R$ with x satisfying (1.829). In other words we see that

$$M_R^* M_R = \frac{2\, f_2\, \Lambda^2}{f_0}\, \frac{k_R^* k_R\, \mathrm{Tr}(k_R^* k_R)}{\mathrm{Tr}((k_R^* k_R)^2)}. \tag{1.832}$$

□

Notice that the Dirac mass M_ν is of the order of the Fermi energy $v = 2M/g$ of (1.788) and hence much smaller. This relative size of the terms in (1.828) is what allows for a see-saw mechanism to take place.

To see this more in detail, we simplify the model to the case of $N = 1$ generations. In this case, the eigenvalues of the matrix (1.828) are of the form

$$\frac{1}{2}\left(\pm M_R \pm \sqrt{M_R^2 + 4v^2}\right), \tag{1.833}$$

where v is the Higgs vacuum (1.788) and M_R is here just a scalar (since $N = 1$), which is of the order of Λ by the result of Lemma 1.226, see (1.832).

The formula (1.833) means that we have two eigenvalues very close to $\pm M_R$ and two others very close to $\pm \frac{v^2}{M_R}$, while, for $N = 1$ generation, the determinant of the matrix (1.828) is equal to $|M_\nu|^4 \sim v^4$.

The see-saw mechanism in particular implies that, while the neutrino masses remain small, they can nonetheless have a large Yukawa coupling. Typical estimates for the large masses of the right-handed neutrinos i.e. the eigenvalues $(m_R)_\sigma$ of M_R are given (cf. [**233**]) by

$$(m_R)_1 \geq 10^7 GeV\,, \quad (m_R)_2 \geq 10^{12} GeV\,, \quad (m_R)_3 \geq 10^{16} GeV\,. \tag{1.834}$$

The case that is especially important for explicit calculations is the τ neutrino, which will have the largest corresponding Yukawa coupling term, of the same order as that of the top quark. This introduces important corrections to the minimal Standard

Model calculations where one would normally assume that the top quark Yukawa coupling is the dominating term.

17.6. The mass relation and the top quark mass.

We now check that the mass relation (1.825) at unification is compatible with the values at ordinary energies.

One comment about notation: there are two conventions in the physics literature on the normalization factor for the Higgs field.

(1) In Veltman [**285**] the kinetic term has a factor of $\frac{1}{2}$
(2) In Mohapatra–Pal it has a factor of 1 (cf. [**233**] equation (1.43))

One passes from one to the other by

$$\varphi_{\mathrm{mp}} = \frac{1}{\sqrt{2}}\,\varphi_{\mathrm{velt}} \tag{1.835}$$

We adopt the second convention (which is the same convention used in [**45**]). Thus, we define the Yukawa couplings in the following way.

DEFINITION 1.227. *For $v = \frac{2M}{g}$ we define the matrices $y_x = (y_x^\sigma)$ with σ the generation index and $x \in \{u, d, \nu, e\}$ by the relation*

$$\frac{v}{\sqrt{2}} y_x^\sigma = m_x^\sigma, \tag{1.836}$$

where the m_x^σ are defined as in (1.826).

The mass of the top quark is governed by the top quark Yukawa coupling $y_t = y_u^\sigma(t)$ with $\sigma = 3$ by the equation

$$m_{\mathrm{top}}(t) = \frac{1}{\sqrt{2}}\frac{2M}{g}\,y_t = \frac{1}{\sqrt{2}}\,v\,y_t, \tag{1.837}$$

where $v = \frac{2M}{g}$ is the vacuum expectation value of the Higgs field. The running of the top quark Yukawa coupling $y_t = y_u^\sigma(t)$, with $\sigma = 3$, is governed by equation (1.843).

As in [**4**], we denote by $Y_2(S)$ the combination

$$Y_2(S) = \sum_\sigma (y_\nu^\sigma)^2 + (y_e^\sigma)^2 + 3\,(y_u^\sigma)^2 + 3\,(y_d^\sigma)^2 \tag{1.838}$$

In these terms the mass constraint (1.825) reads as

$$Y_2(S) = 4\,g^2, \tag{1.839}$$

i.e. as

$$\frac{v^2}{2}\sum_\sigma (y_\nu^\sigma)^2 + (y_e^\sigma)^2 + 3\,(y_u^\sigma)^2 + 3\,(y_d^\sigma)^2 = \,2\,g^2\,v^2, \tag{1.840}$$

with $v = \frac{2M}{g}$ the vacuum expectation value of the Higgs, as above.

Assume that the relation (1.840) holds at a unification scale Λ of 10^{17} GeV, neglecting the discrepancy of Figure 33 and assuming the exact relation (1.783) between the coupling constants.

In the first approximation, as is usually done in the context of the minimal Standard Model, one can choose to neglect all the Yukawa couplings other than that of the

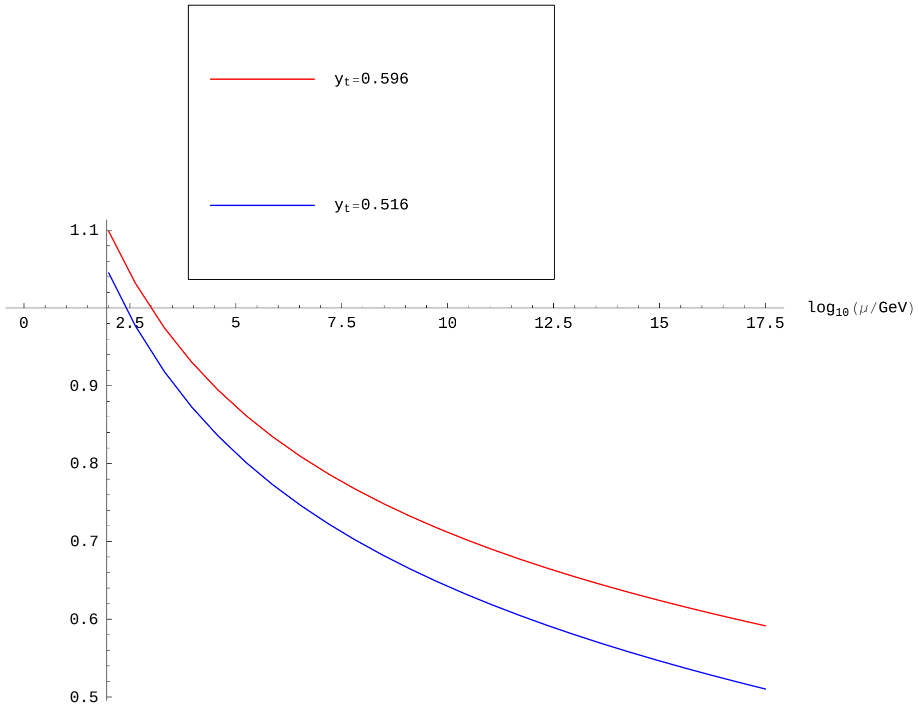


FIGURE 34. The running of the top quark Yukawa coupling.

top quark, that is, y_u^σ, with $\sigma = 3$. This yields an approximate form of (1.825) given by

$$y_u^\sigma = \frac{2}{\sqrt{3}}\, g, \qquad \text{for} \quad \sigma = 3. \tag{1.841}$$

We consider the Yukawa couplings $(y_\cdot^\sigma)$ as depending on the energy scale through their renormalization group equation (cf. [**4**], [**43**], [**243**]). We set

$$t = \log(\frac{\Lambda}{M_Z}) \quad \text{and} \quad \mu = M_Z e^t. \tag{1.842}$$

We consider in particular the top quark case $y_u^\sigma(t)$ for $\sigma = 3$. The running of the top quark Yukawa coupling $y_t = y_u^\sigma(t)$, with $\sigma = 3$, is governed by the differential equation (cf. [**267**] equation (2.143) and equation (A9) of [**4**])

$$\frac{dy_t}{dt} = \frac{1}{16\pi^2}\left(\frac{9}{2} y_t^3 - \left(a\, g_1^2 + b\, g_2^2 + c\, g_3^2\right) y_t\right), \tag{1.843}$$

with

$$(a, b, c) = (\frac{17}{12}, \frac{9}{4}, 8).$$

The value of g at a unification scale of 10^{17} GeV is ~ 0.517. Thus, (1.841) gives the simplified relation

$$y_t = \frac{2}{\sqrt{3}} g \sim 0.597\,, \qquad \text{for} \quad t \sim 34.6\,. \tag{1.844}$$

Thus, in first approximation, numerical integration of the differential equation (1.843) with the boundary condition (1.844) gives the value at ordinary energies $y_0 \sim 1.102$, which means a top quark mass of the order of $\frac{1}{\sqrt{2}} y_0\, v \sim 173.683\, y_0$ GeV. This is not in good agreement with the physical value. However, our assumption that the top quark Yukawa coupling is the only dominant term in (1.840) would only be justified in the minimal Standard Model. In our model with neutrino mixing and the Majorana mass term, the see-saw mechanism discussed in §17.5 above implies that the Yukawa coupling term for the τ neutrino is also of the same order of magnitude as that of the top quark, hence it introduces a visible correction to (1.844)
In fact, although the known upper bound on the mass of the τ neutrino is only of the order of (cf. [**233**])

$$m_{\nu_\tau} \leq 18.2 \text{ MeV},$$

which corresponds to the small eigenvalue of (1.833), the see-saw mechanism allows for a large Yukawa coupling term y_ν^σ, with $\sigma = 3$ by the relation (1.833) and (1.834), which turns out to be comparable to that of the top quark. We can then make the simplifying assumption that, at unification, the Yukawa coupling y_ν^σ, with $\sigma = 3$ for the τ neutrino to be the same as the y_u^σ, with $\sigma = 3$ of the top quark.
When we take into account the contribution of the Yukawa coupling of the τ neutrino in this way, this introduces in (1.844) a correction factor of $\sqrt{3/4}$. In fact, for $x_t = y_\nu^\sigma(t)$ and $y_t = y_u^\sigma(t)$, with $\sigma = 3$, we now have

$$Y_2(S) \sim x_t^2 + 3y_t^2 \sim \frac{4}{3} \cdot 3y_t^2 = 4\, y_t^2 \Rightarrow y_t \sim g \tag{1.845}$$

This replaces the initial value (1.844) of y_t by $y_t \sim 0.516$ and has (Figure 34) the effect of lowering the value of y_0 to $y_0 \sim 1.04$, which yields an acceptable value for the top quark mass.

17.7. The self-interaction of the gauge bosons.

We now consider the self-interaction terms $\mathcal{L}_g$ for the gauge fields, as in (1.445) and we show that they can be written as a sum

$$\mathcal{L}_g = \mathcal{L}_g^{mass} + \mathcal{L}_g^{YM} + \mathcal{L}_{feyn} \tag{1.846}$$

of terms of the following form.

(1) $\mathcal{L}_g^{mass}$ consists of the mass terms for the $W^\pm$ and the Z^0.
(2) $\mathcal{L}_g^{YM}$ consists of the Yang-Mills interaction $-\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu}$ for the gauge fields B_μ, W_μ^a, g_μ^a.
(3) $\mathcal{L}_{feyn}$ are the Feynman gauge fixing terms for all the gauge fields.

LEMMA 1.228. *One has*

$$\mathcal{L}_g = -M^2 W_\mu^+ W_\mu^- - \frac{1}{2c_w^2} M^2 Z_\mu^0 Z_\mu^0 - \frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} - \frac{1}{2} \sum_a (\sum_\mu \partial_\mu G_\mu^a)^2. \tag{1.847}$$

PROOF. It is enough to show that the expression

$$-\partial_\nu W^+_\mu \partial_\nu W^-_\mu - \tfrac{1}{2}\partial_\nu(c_w\, Z^0_\mu + s_w\, A_\mu)\partial_\nu(c_w\, Z^0_\mu + s_w\, A_\mu) - igc_w(\partial_\nu Z^0_\mu(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - Z^0_\nu(W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + Z^0_\mu(W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) - igs_w(\partial_\nu A_\mu(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - A_\nu(W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + A_\mu(W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) - \tfrac{1}{2}g^2 W^+_\mu W^-_\mu W^+_\nu W^-_\nu + \tfrac{1}{2}g^2 W^+_\mu W^-_\nu W^+_\mu W^-_\nu + g^2 c_w^2(Z^0_\mu W^+_\mu Z^0_\nu W^-_\nu - Z^0_\mu Z^0_\mu W^+_\nu W^-_\nu) + g^2 s_w^2(A_\mu W^+_\mu A_\nu W^-_\nu - A_\mu A_\mu W^+_\nu W^-_\nu) + g^2 s_w c_w(A_\mu Z^0_\nu(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - 2A_\mu Z^0_\mu W^+_\nu W^-_\nu)$$

coincides with the Yang–Mills action of the SU(2) gauge field.
In fact, the kinetic terms will then combine with those of the B-field, namely

$$-\frac{1}{2}\partial_\nu(-s_w\, Z^0_\mu + c_w\, A_\mu)\partial_\nu(-s_w\, Z^0_\mu + c_w\, A_\mu).$$

One can rewrite the expression above in terms of $W_{\mu 3} = s_w\, A_\mu + c_w\, Z^0_\mu$. This gives

$$-\partial_\nu W^+_\mu \partial_\nu W^-_\mu - \tfrac{1}{2}\partial_\nu W_{\mu 3}\partial_\nu W_{\mu 3} - ig(\partial_\nu W_{\mu 3}(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - W_{\nu 3}(W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + W_{\mu 3}(W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) - \tfrac{1}{2}g^2 W^+_\mu W^-_\mu W^+_\nu W^-_\nu + \tfrac{1}{2}g^2 W^+_\mu W^-_\nu W^+_\mu W^-_\nu + g^2(W_{\mu 3} W^+_\mu W_{\nu 3} W^-_\nu - W_{\mu 3} W_{\mu 3} W^+_\nu W^-_\nu).$$

Using $W^+_\mu = \frac{W_{\mu 1} - iW_{\mu 2}}{\sqrt{2}}$ and $W^-_\mu = \frac{W_{\mu 1} + iW_{\mu 2}}{\sqrt{2}}$, one checks that it coincides with the Yang-Mills action functional $-\frac{1}{4}\, F^a_{\mu\nu}\, F^{\mu\nu}_a$ of the SU(2) gauge field $W_{\mu j}$.
More precisely, let

$$\nabla_\mu = \partial_\mu - i\frac{g}{2}W^\mu_a \sigma_a.$$

One then has

$$[\nabla_\mu, \nabla_\nu] = -i\frac{g}{2}(\partial_\mu W^\nu_a - \partial_\nu W^\mu_a)\sigma_a + (-i\frac{g}{2})^2(W^\mu_b W^\nu_c \sigma_b\, \sigma_c - W^\nu_c W^\mu_b \sigma_c\, \sigma_b).$$

Using $\sigma_b\, \sigma_c - \sigma_c\, \sigma_b = 2i\, \epsilon_{abc}\, \sigma_a$, this gives

(1.848) $$F^a_{\mu\nu} = \partial_\mu W^\nu_a - \partial_\nu W^\mu_a + g\, \epsilon_{abc}\, W^\mu_b W^\nu_c.$$

One then checks directly that the above expression coincides with

(1.849) $$-\frac{1}{4}\, F^a_{\mu\nu}\, F^{\mu\nu}_a - \frac{1}{2}\sum_a(\sum_\mu \partial_\mu W^a_\mu)^2.$$

Notice that the addition of the Feynman gauge fixing term $-\frac{1}{2}(\sum_\mu \partial_\mu G^\mu)^2$ to the kinetic term $-\frac{1}{4}\, |dG|^2$ of the Yang-Mills action for each of the gauge fields G^μ gives kinetic terms of the form $-\frac{1}{2}\, \partial_\nu G^\mu\, \partial_\nu G^\mu$ and very simple propagators. □

COROLLARY 1.229. *The terms $\mathcal{L}_g^{YM}$ of $\mathcal{L}_g$ are obtained from the asymptotic expansion of the spectral action functional* (1.772)*, under the condition* (1.783) *on the coupling constants.*

PROOF. The part $\mathcal{L}_g^{YM}$ of (1.847) is of the form $-\frac{1}{4}\, F^a_{\mu\nu}\, F^{\mu\nu}_a$, for the various gauge fields B_μ, W^a_μ, and g^a_μ. This is shown in Lemma 1.228 for the more complicated case of the SU(2) gauge field. The factor 1/4 in front of this Yang–Mills terms shows that it can be identified with the corresponding Yang–Mills terms in the asymptotic expansion (1.754) of the spectral action of Theorem 1.212 under the condition (1.783) on the parameters, which gives the form (1.784) of the asymptotic expansion of the spectral action and imposes the relation among the coupling constants at unification, as we discussed in §17.1 above. □

Notice that, while the Yang–Mills terms are obtained from the spectral action, as we have seen in Theorem 1.212, the gauge fixing term is not obtained from the spectral action. It is added separately, together with the ghost fields. The mass terms $\mathcal{L}_g^{mass}$ are discussed in the next section in combination with the minimal coupling terms of the Higgs.

17.8. The minimal coupling of the Higgs field.

We now consider the terms $\mathcal{L}_{Hg}$ of the Lagrangian, as in (1.446), which contain the minimal coupling terms of the Higgs fields with the gauge fields. We add to $\mathcal{L}_{Hg}$ the mass terms

$$\mathcal{L}_g^{mass} = -M^2 W_\mu^+ W_\mu^- - \frac{1}{2c_w^2} M^2 Z_\mu^0 Z_\mu^0 \tag{1.850}$$

of $\mathcal{L}_g$ as in (1.847). This is, by construction, a sum of terms labeled by μ. Each of them contains three kinds of terms, according to the number of derivatives.

The following result compares this expression with the minimal coupling terms which we get from the spectral action.

LEMMA 1.230. *With the notation* (1.798) *of Lemma 1.220, the minimal coupling terms* (1.446) *are given by*

$$\mathcal{L}_{Hg} + \mathcal{L}_g^{mass} = -\frac{1}{2} |D_\mu \varphi|^2, \tag{1.851}$$

with $D_\mu \varphi$ given by (1.747), *with $g_2 = g$, $g_1 = g'$.*

PROOF. We have from (1.747)

$$D_\mu \varphi = \partial_\mu \varphi + \frac{i}{2} g W_\mu^\alpha \varphi \sigma^\alpha - \frac{i}{2} g' B_\mu \varphi, \quad g' = \tan(\theta_w) g \tag{1.852}$$

where, by Lemma 1.220, we have

$$\begin{aligned} \varphi &= (\varphi_1, \varphi_2) = (\tfrac{2M}{g} + H - i\phi^0, -i\sqrt{2}\,\phi^+), \\ B_\mu &= c_w A_\mu - s_w Z_\mu^0, \\ W_{\mu 3} &= s_w A_\mu + c_w Z_\mu^0. \end{aligned}$$

The matrix $W_\mu^\alpha \sigma^\alpha$ in (1.852) is of the form

$$\begin{aligned} W_\mu^\alpha \sigma^\alpha &= \begin{pmatrix} s_w A_\mu + c_w Z_\mu^0 & W_\mu^1 - i W_\mu^2 \\ W_\mu^1 + i W_\mu^2 & -s_w A_\mu - c_w Z_\mu^0 \end{pmatrix} \\ &= \begin{pmatrix} s_w A_\mu + c_w Z_\mu^0 & \sqrt{2} W_\mu^+ \\ \sqrt{2} W_\mu^- & -s_w A_\mu - c_w Z_\mu^0 \end{pmatrix}. \end{aligned}$$

The kinetic terms in $\mathcal{L}_{Hg}$ are simply

$$-\frac{1}{2}\partial_\mu H \partial_\mu H - \partial_\mu \phi^+ \partial_\mu \phi^- - \frac{1}{2}\partial_\mu \phi^0 \partial_\mu \phi^0$$

and one checks directly that these are indeed obtained from the expression on the right-hand side of (1.851), using (1.852).

Let us consider then the terms with no derivatives. The combination $W_\mu^\alpha \varphi \sigma^\alpha$ is given by

$$((W_\mu^\alpha \varphi \sigma^\alpha)_1, (W_\mu^\alpha \varphi \sigma^\alpha)_2),$$

with

$$(W_\mu^\alpha\, \varphi\, \sigma^\alpha)_1 = (\tfrac{2M}{g} + H - i\phi^0)(s_w\, A_\mu + \, c_w\, Z_\mu^0) - 2i\, \phi^+\, W_\mu^-,$$

$$(W_\mu^\alpha\, \varphi\, \sigma^\alpha)_2 = (\tfrac{2M}{g} + H - i\phi^0)\sqrt{2}\, W_\mu^+ + i\, \sqrt{2}\, \phi^+(s_w\, A_\mu + \, c_w\, Z_\mu^0)).$$

The term $B_\mu\, \varphi$ is given by

$$B_\mu\, \varphi = \left((\frac{2M}{g} + H - i\phi^0)(c_w\, A_\mu - s_w\, Z_\mu^0), - i\, \sqrt{2}\, \phi^+(c_w\, A_\mu - s_w\, Z_\mu^0) \right).$$

Notice that, for physical reasons, the terms in MA_μ have to disappear, since the presence of such a term would imply a massive photon. The terms do indeed cancel, because of the relation $g' = \tan(\theta_w)\, g$. In fact, in the expression

$$+\frac{i}{2}\, g\, W_\mu^\alpha\, \varphi\, \sigma^\alpha - \frac{i}{2}\, g'\, B_\mu\, \varphi$$

in (1.852), the terms in MA_μ appear in the combination

$$i(s_w - \frac{g'}{g} c_w) M A_\mu \equiv 0,$$

where we use the fact that $g'/g = s_w/c_w$.

Thus, we obtain

$$W_\mu^\alpha\, \varphi\, \sigma^\alpha - \tan(\theta_w)\, B_\mu\, \varphi = (X_1, X_2), \tag{1.853}$$

where we have

$$X_1 = (\tfrac{2M}{g} + H - i\phi^0)\tfrac{1}{c_w}\, Z_\mu^0 - 2i\, \phi^+\, W_\mu^-$$

$$X_2 = (\tfrac{2M}{g} + H - i\phi^0)\sqrt{2}\, W_\mu^+ + i\, \sqrt{2}\, \phi^+(2\, s_w\, A_\mu + \, (c_w - \tfrac{s_w^2}{c_w})\, Z_\mu^0).$$

One has to multiply by $\frac{g}{2}\sqrt{-1}$ and then take $-\frac{1}{2}$ of the norm squared. A direct computation gives

$$\begin{gathered}
-M^2 W_\mu^+ W_\mu^- - \tfrac{1}{2c_w^2} M^2 Z_\mu^0 Z_\mu^0 - gMW_\mu^+ W_\mu^- H - \tfrac{1}{2} g \tfrac{M}{c_w^2} Z_\mu^0 Z_\mu^0 H - ig \tfrac{s_w^2}{c_w} M Z_\mu^0 (W_\mu^+ \phi^- - \\
W_\mu^- \phi^+) + igs_w M A_\mu (W_\mu^+ \phi^- - W_\mu^- \phi^+) - \tfrac{1}{4} g^2 W_\mu^+ W_\mu^- \left(H^2 + (\phi^0)^2 + 2\phi^+\phi^-\right) - \\
\tfrac{1}{8} g^2 \tfrac{1}{c_w^2} Z_\mu^0 Z_\mu^0 \left(H^2 + (\phi^0)^2 + 2(2s_w^2 - 1)^2 \phi^+\phi^-\right) - \tfrac{1}{2} g^2 \tfrac{s_w^2}{c_w} Z_\mu^0 \phi^0 (W_\mu^+ \phi^- + W_\mu^- \phi^+) - \\
\tfrac{1}{2} i g^2 \tfrac{s_w^2}{c_w} Z_\mu^0 H (W_\mu^+ \phi^- - W_\mu^- \phi^+) + \tfrac{1}{2} g^2 s_w A_\mu \phi^0 (W_\mu^+ \phi^- + W_\mu^- \phi^+) + \\
\tfrac{1}{2} i g^2 s_w A_\mu H (W_\mu^+ \phi^- - W_\mu^- \phi^+) - g^2 \tfrac{s_w}{c_w} (1 - 2s_w^2) Z_\mu^0 A_\mu \phi^+ \phi^- - g^2 s_w^2 A_\mu A_\mu \phi^+ \phi^-.
\end{gathered}$$

As can be seen from (1.446), the terms with no derivatives in $\mathcal{L}_{Hg} + \mathcal{L}_g^{mass}$ are exactly the same as those listed here above, except for the replacement of the term $(2c_w^2 - 1)$ by $(1 - 2s_w^2)$ in the second to last term, which has no effect since $s_w^2 + c_w^2 = 1$.

We now need to take care of the terms with one derivative. With the notation as in (1.853), we compute the cross terms of

$$-\frac{1}{2}\, |(\partial_\mu \varphi_1, \partial_\mu \varphi_2) + \frac{ig}{2}\, (X_1, X_2)|^2,$$

i.e. the terms

$$-\frac{1}{2}\, (\partial_\mu \bar\varphi_1\, \frac{ig}{2} X_1 - \partial_\mu \varphi_1\, \frac{ig}{2} \bar X_1 + \partial_\mu \bar\varphi_2\, \frac{ig}{2} X_2 - \partial_\mu \varphi_2\, \frac{ig}{2} \bar X_2).$$

The computation gives

$$-\tfrac{1}{2}ig\left(W_\mu^+(\phi^0\partial_\mu\phi^- - \phi^-\partial_\mu\phi^0) - W_\mu^-(\phi^0\partial_\mu\phi^+ - \phi^+\partial_\mu\phi^0)\right) + \tfrac{1}{2}g\left(W_\mu^+(H\partial_\mu\phi^- - \phi^-\partial_\mu H) + W_\mu^-(H\partial_\mu\phi^+ - \phi^+\partial_\mu H)\right) + \tfrac{1}{2}g\tfrac{1}{c_w}Z_\mu^0(H\partial_\mu\phi^0 - \phi^0\partial_\mu H) - ig\tfrac{1-2c_w^2}{2c_w}Z_\mu^0(\phi^+\partial_\mu\phi^- - \phi^-\partial_\mu\phi^+) + igs_wA_\mu(\phi^+\partial_\mu\phi^- - \phi^-\partial_\mu\phi^+) + M\left(\tfrac{1}{c_w}Z_\mu^0\partial_\mu\phi^0 + W_\mu^+\partial_\mu\phi^- + W_\mu^-\partial_\mu\phi^+\right),$$

which agrees with the sum of terms with one derivative in $\mathcal{L}_{Hg}$ as in (1.446). $\square$

17.9. The Higgs field self-interaction.

We now turn to the Higgs self-coupling terms $\mathcal{L}_H$ of the Standard Model, which are of the form (1.444). Under the natural assumption (1.854) below (a standard relation giving the Higgs mass, cf. [**285**]), we obtain the following description of the terms $\mathcal{L}_H$.

LEMMA 1.231. *Let φ be given by* (1.810) *and assume that*

$$\alpha_h = \frac{m_h^2}{4\,M^2}. \tag{1.854}$$

Then one has

$$\mathcal{L}_H = -\frac{1}{8}g^2\alpha_h\,|\varphi|^4 + (\alpha_h\,M^2\,-\frac{\beta_h}{2})\,|\varphi|^2. \tag{1.855}$$

with m_h, α_h and β_h as in (1.444).

PROOF. The expression (1.444) can be simplified in terms of the field $\psi = (\psi_1, \psi_2)$ of (1.808) and (1.809). The quartic term is simply given by

$$-\frac{1}{8}g^2\alpha_h\left(H^4 + (\phi^0)^4 + 4(\phi^+\phi^-)^2 + 4(\phi^0)^2\phi^+\phi^- + 4H^2\phi^+\phi^- + 2(\phi^0)^2H^2\right)$$
$$= -\frac{1}{8}g^2\alpha_h|\psi|^4,$$

since $|\psi|^2 = |\psi_1|^2 + |\psi_2|^2 = H^2 + (\phi^0)^2 + 2\,\phi^+\phi^-$.
The cubic term is

$$-g\alpha_h M\left(H^3 + H\phi^0\phi^0 + 2H\phi^+\phi^-\right) = -g\alpha_h M\,H\,|\psi|^2,$$

which arises in the expansion of

$$-\frac{1}{8}g^2\alpha_h\,|\varphi|^4, \tag{1.856}$$

with φ given by (1.810), so that

$$|\varphi|^2 = |\psi|^2 + \frac{4M}{g}\,H + \frac{4M^2}{g^2}$$

and

$$|\varphi|^4 = |\psi|^4 + \frac{8M}{g}\,H\,|\psi|^2 + \frac{16M^2}{g^2}\,H^2 + \frac{8M^2}{g^2}\,|\psi|^2 + \frac{16M^4}{g^4} + \frac{32M^3}{g^3}\,H\,.$$

Thus, the natural invariant expression with no tadpole (i.e. with the expansion in H at an extremum) is

$$-\frac{1}{8}g^2\alpha_h\,|\varphi|^4 + \,\alpha_h\,M^2\,|\varphi|^2. \tag{1.857}$$

It expands as

$$-\frac{1}{8}g^2\alpha_h|\psi|^4 - g\alpha_h M\, H\,|\psi|^2 - 2\alpha_h\, M^2\, H^2 + \frac{2M^4}{g^2}\alpha_h, \tag{1.858}$$

which takes care of the constant term $+\frac{2M^4}{g^2}\alpha_h$ in (1.444). Thus, we get

$$\mathcal{L}_H = (-\frac{1}{8}g^2\alpha_h\,|\varphi|^4 + \alpha_h\,M^2\,|\varphi|^2) + (2\alpha_h\,M^2 - \frac{1}{2}m_h^2)H^2 - \frac{\beta_h}{2}\,|\varphi|^2, \tag{1.859}$$

since the quadratic "tadpole" term in (1.444) is

$$-\beta_h\left(\frac{2M^2}{g^2} + \frac{2M}{g}H + \frac{1}{2}(H^2 + \phi^0\phi^0 + 2\phi^+\phi^-)\right) = -\frac{\beta_h}{2}\,|\varphi|^2. \tag{1.860}$$

The assumption (1.854) of the lemma implies that the coefficient of the term in H^2 in (1.859) vanishes. □

This completes the checking of all the terms of $\mathcal{L}_{SM}$ and their relation to the spectral action functional. Thus, this completes the proof of Theorem 1.217. □

17.10. The Higgs scattering parameter and the Higgs mass.

When written in terms of $\mathbf{H}$, and using (1.783), the quartic term

$$\frac{f_0}{2\,\pi^2}\int b\,|\varphi|^4\,\sqrt{g}\,d^4x = \frac{\pi^2}{2\,f_0}\frac{b}{a^2}\int |\mathbf{H}|^4\,\sqrt{g}\,d^4x$$

gives a further relation in our theory between the $\lambda|\mathbf{H}|^4$ coupling in (1.784) and the gauge couplings to be imposed at the scale Λ. This is of the form

$$\tilde\lambda(\Lambda) = g_3^2\,\frac{b}{a^2}, \tag{1.861}$$

where we have

$$\lambda(\mu) = 4\tilde\lambda(\mu), \qquad \text{with } \mu = M_Z e^t \tag{1.862}$$

as in (1.842). The reason for passing from λ to $\tilde\lambda$ is the change of normalization (1.835).

As above, let y_x^σ, with $\sigma = 1, 2, 3$ the generation index and $x \in \{u, d, \nu, e\}$, be defined as in (1.836), and let $\mu = M_Z e^t$, with $t = \log(\frac{\Lambda}{M_Z})$ as in (1.842). The running of the top quark Yukawa coupling ($x = u$, $\sigma = 3$) is given as in (1.843) above.

If we assume, in first approximation, that the top quark Yukawa coupling is much larger than all the other Yukawa couplings, the relation (1.861) simplifies. One gets $a \sim 3\,m_{\rm top}^2$ and $b \sim 3\,m_{\rm top}^4$, where $m_{\rm top} = m_u^\sigma$, with $\sigma = 3$ in the notation of (1.826), so that

$$\tilde\lambda(\Lambda) \sim \frac{4}{3}\,\pi\,\alpha_3(\Lambda)\,. \tag{1.863}$$

In terms of the Higgs scattering parameter α_h of the Standard Model, (1.863) reads

$$\alpha_h(\Lambda) \sim \frac{8}{3} \tag{1.864}$$

which agrees with [**187**] equation (1). Therefore, the value of $\lambda = 4\,\tilde\lambda$ at the unification scale of $\Lambda = 10^{17}$ GeV is $\lambda_0 \sim 0.356$, showing that one does not go outside the perturbation domain.

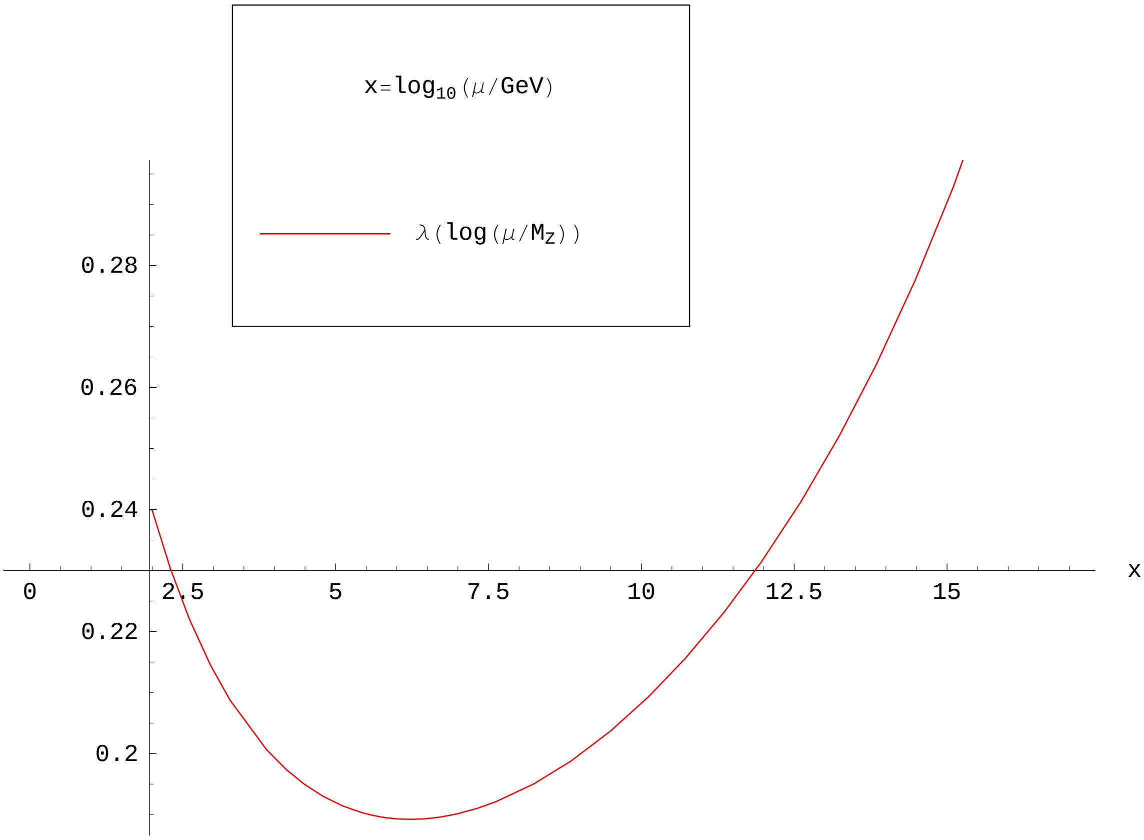


FIGURE 35. The running of the quartic Higgs coupling.

Equation (1.863) can be used, together with the RG equations for λ and $y_u^\sigma(t)$, with $\sigma = 3$, to determine the Higgs mass at the low-energy scale M_Z.
For simplicity of notation, in the following we write

$$y_t = y_u^\sigma(t), \quad \text{with} \quad \sigma = 3 \tag{1.865}$$

for the Yukawa coupling of the top quark as a function of t related to the energy scale by $\mu = M_Z e^t$.
We have (cf. [**267**] equations (2.141), (2.142), (4.2) and the formula (A15) of [**4**]) the equation

$$\frac{d\lambda}{dt} = \lambda\gamma + \frac{1}{8\pi^2}(12\lambda^2 + B) \tag{1.866}$$

where

$$\begin{aligned} \gamma &= \frac{1}{16\pi^2}(12y_t^2 - 9g_2^2 - 3g_1^2) \\ B &= \frac{3}{16}(3g_2^4 + 2g_1^2\, g_2^2 + g_1^4) - 3y_t^4 \,. \end{aligned} \tag{1.867}$$

The Higgs mass is then given by

$$m_H^2 = 8\lambda \, \frac{M^2}{g^2} \,, \quad m_H = \sqrt{2\lambda}\, \frac{2M}{g} \tag{1.868}$$

The numerical solution to these equations with the boundary value $\lambda_0 = 0.356$ at $\Lambda = 10^{17}$ GeV gives $\lambda(M_Z) \sim 0.241$ and a Higgs mass of the order of 170 GeV.
Once again, we need to correct for the fact that, in our model with neutrino mixing and Majorana mass terms for the neutrinos, the see-saw mechanism of §17.5 introduces a large Yukawa coupling for the τ neutrino, which is of the same order as that of the top quark. Thus, as we did in §17.6, we need to correct this first approximation by a factor that accounts for the τ neutrino Yukawa coupling. Thus, the factor of 4/3 in (1.863) should be corrected to 1 as in (1.845) below. One can check by direct calculation that this does not affect substantially the estimate we obtain for the Higgs mass which is then around 168 GeV.

Thus, we see that this model predicts a heavy Higgs at around 168 GeV. For a similar analysis of the Higgs mass in other variants of the noncommutative geometry approach to the Standard Model, see [**41**] and [**187**].

17.11. The gravitational terms.

We have seen in Theorem 1.217 above how to recover the Standard Model Lagrangian $\mathcal{L}_{SM}$ from the spectral action. We have, for simplicity, considered the case of flat space, since by covariance this suffices to show we can recover $\mathcal{L}_{SM}$ in general. We now discuss separately the gravitational terms that are obtained from the spectral action and we compare them with the formulation of higher derivatives gravity as an effective field theory that we recalled briefly in §9.8 above.

As we have seen in Theorem 1.212 above, the gravitational terms in the spectral action are of the form

$$\int \left(\frac{1}{2\kappa_0^2}\, R + \alpha_0 \, C_{\mu\nu\rho\sigma} \, C^{\mu\nu\rho\sigma} + \gamma_0 + \tau_0 \, R^* R^* - \xi_0 \, R \, |\mathbf{H}|^2 \right) \sqrt{g} \, d^4 x. \tag{1.869}$$

Notice that, with respect to the general form of the higher derivatives gravity with quadratic terms in the curvature of (1.463), in (1.869) we do not have the term in R^2, but only the term quadratic in the Weyl curvature and the topological term $E = R^* R^*$ with

$$\frac{1}{32\pi^2} \int R^* R^* \sqrt{g} \, d^4 x = \chi(M). \tag{1.870}$$

The conceptual explanation of the absence of the R^2 terms in (1.869) can be found in [**47**] where the conformal invariance of the spectral action is analyzed.

The running of the coefficients of the Euclidean higher derivative terms in (1.463), determined by the renormalization group equation, is gauge independent and is given by (see e.g. [**11**], equations (4.49) and (4.71). See also [**122**], [**58**])

$$\begin{aligned}
\beta_\eta &= -\frac{1}{(4\pi)^2} \, \frac{133}{10} \eta^2 \\
\beta_\omega &= -\frac{1}{(4\pi)^2} \, \frac{25 + 1098\, \omega + 200\, \omega^2}{60} \eta \\
\beta_\theta &= \frac{1}{(4\pi)^2} \frac{7(56 - 171\, \theta)}{90} \eta.
\end{aligned}$$

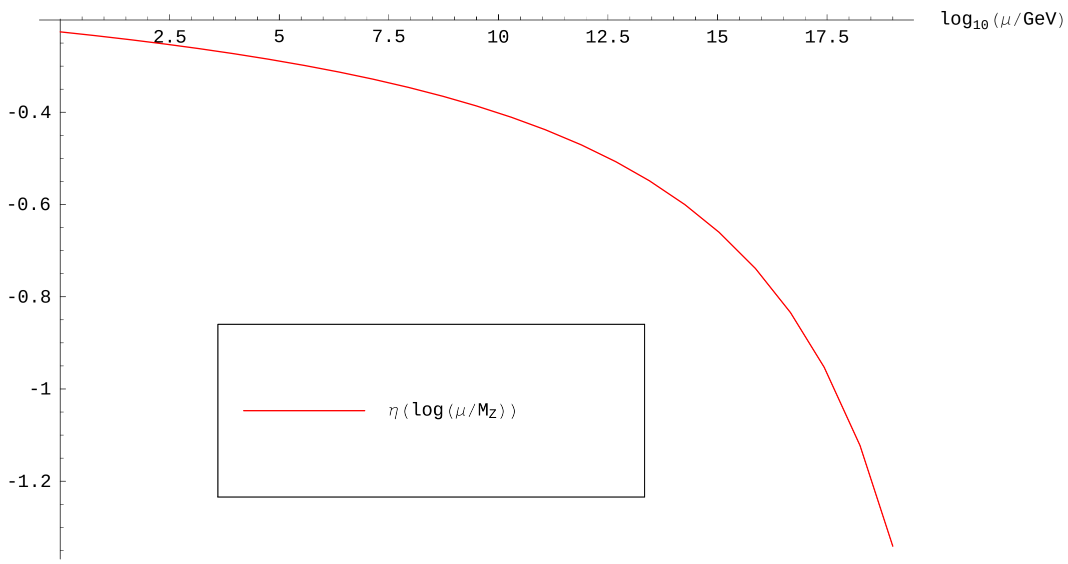


FIGURE 36. The running of the Weyl curvature term in (1.463).

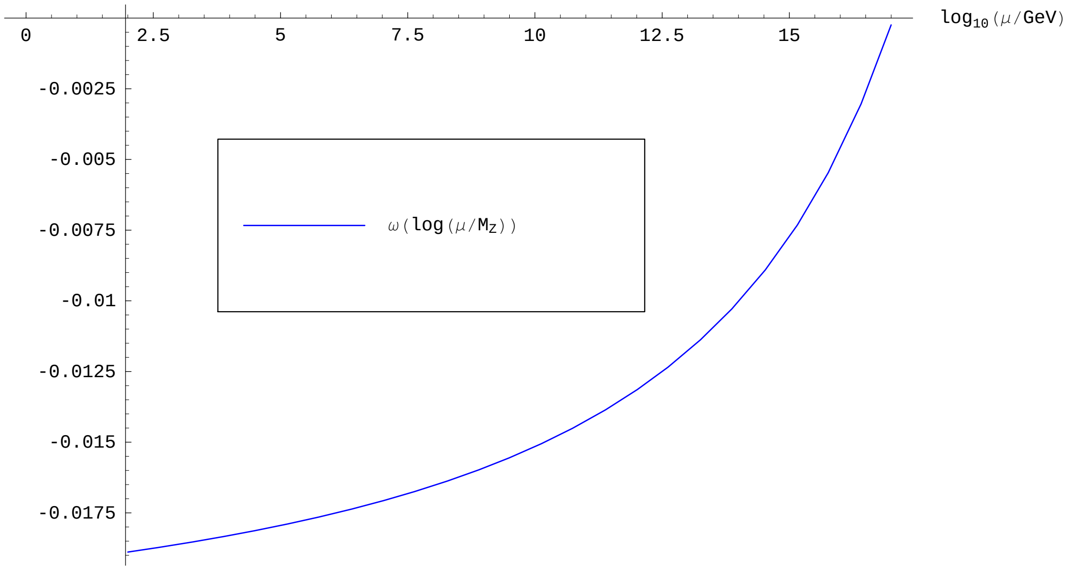


FIGURE 37. The running of the ratio of the coefficients of the R^2 term and the Weyl term in (1.463).

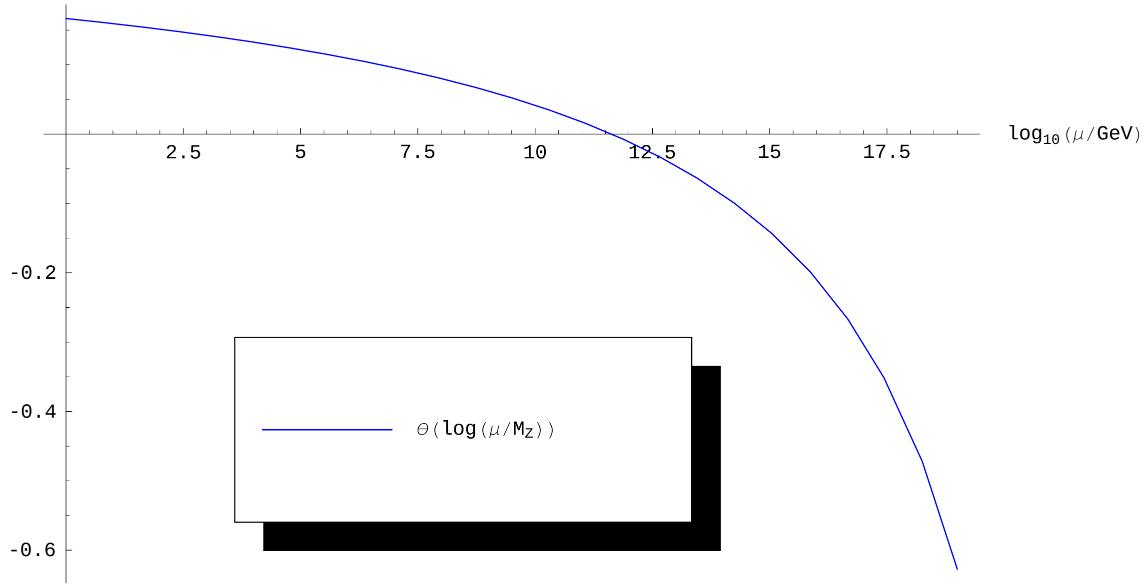


FIGURE 38. The running of the ratio of the coefficients of the topological term and the Weyl term in (1.463).

The graphs are shown in Figures 36, 37, 38. Notice that the infrared behavior of these terms approaches the fixed point $\eta = 0$, $\omega = -0.0228$, $\theta = 0.327$. The coefficient η goes to zero in the infrared limit, sufficiently slowly, so that, up to scales of the order of the size of the universe, its inverse remains $O(1)$. On the other hand, $\eta(t)$, $\omega(t)$ and $\theta(t)$ have a common singularity at an energy scale of the order of 10^{23} GeV, which is above the Planck scale. Moreover, within the energy scales that are of interest to our model $\eta(t)$ is neither too small nor too large (it does not vary by more than a single order of magnitude between the Planck scale and infrared energies). The only known experimental constraints on the values of the coefficients of the quadratic curvature terms $R_{\mu\nu}R^{\mu\nu}$ and R^2 at low energy are very weak and predict that their value should not exceed 10^{74} (cf. e.g. [**122**]). In our case, this is guaranteed by the running described above. Note that we have neglected the coupling $R\,\mathbf{H}^2$ with the Higgs field which ought to be taken into account in a finer analysis.

For a discussion of the Einstein term and its running and of the cosmological term, we refer the reader to [**52**], as well as for a preliminary discussion of the naturalness problem and its relation to the tadpole term. The naturalness problem will be discussed here at the end of Chapter 4, when we draw some comparisons between the physics described in this chapter and the number theoretic content of Chapters 2, 3 and 4.

17.12. The parameters of the Standard Model.

We summarize here the geometric interpretation of the various free parameters that appear in the Standard Model and in the gravitational terms of the action.
The Yukawa coupling parameters consist of 31 real parameters that assign the masses of quarks and leptons, the angles and phases of the Cabibbo–Kobayashi–Maskawa mixing matrix for quarks and the Pontecorvo–Maki–Nakagawa–Sakata mixing matrix for leptons, as well as the Majorana mass parameters for neutrinos. All these parameters describe the choice of a point in the moduli space $\mathcal{C}_\ell \times \mathcal{C}_q$ of Dirac operators for the finite noncommutative geometry F, as described in §13.5 above. Notice that there are interesting loci on this moduli space and interesting compactifications and the question arises of understanding how these may be related to the dynamics introduced by the renormalization group equations.
The coupling constants are related at unification to the parameter f_0 of the asymptotic expansion of the spectral action by (1.783), while the parameters that appear in the gravitational terms are also described in terms of the coefficients f_0, f_2, and f_4 in the expansion of the spectral action, as shown in [**52**].
The model of [**52**] prescribes some relations between these parameters. One is the mass relation at unification described in §17.4 and in §17.6. The Weinberg angle is also obtained, as in the grand unified theories where one has the same phenomenon of the merging of the coupling constants at unification. The Higgs mass is also derived from a relation at unification between the quartic coupling $\lambda|\mathbf{H}|^4$ and the gauge couplings, as described in §17.10. The remaining parameter of the Standard Model described in §9 above, namely the QCD vacuum angle, does not matter in our model, since we are considering a compact 4-manifold M, see [**38**].

18. Functional integral

While renormalization works remarkably well for the quantization of the classical fields involved in the Standard Model, this latter perturbative technique fails dramatically when one tries to deal with the gravitational field $g_{\mu\nu}$. The technical reason for the notorious difficulty of quantizing the $g_{\mu\nu}$ in the traditional perturbative way is the clash with either renormalizability or unitarity. The higher derivative terms such as the ones provided by the spectral action correct the non-renormalizability (cf. [**274**]) but unitarity fails.

In many ways this result is not surprising. Indeed many of the basic notions of the traditional formalism of Quantum Field Theory, such as particles, scattering matrices, etc. rely heavily on the flat geometry of Minkowski space and the related Poincaré symmetry group. Treating the quantization of the $g_{\mu\nu}$ in the same way would – if successful– produce a quantum field theory of the $g_{\mu\nu}$ on Minkowski space: a strange result indeed when viewed from the geometric standpoint!

Thus, rather than treating the gravitational field $g_{\mu\nu}$ as just another field in a given background, one needs a framework which makes it possible to treat all possible background geometries on the same footing and to formulate the Feynman integral including the gravitational degrees of freedom. The formulation of the Feynman integral for gravity is highly dependent upon the chosen geometric set-up. Loosely speaking, in the Euclidean framework one should perform the functional integral on the space of all geometries and the "no boundary proposal" ([**164**]) suggests that one should take all the 4-dimensional geometries with a fixed 3-dimensional boundary as the natural variable in the functional integral computing the wave function evaluated on the given 3-geometry.

The standard Riemannian geometric paradigm contains as basic ingredients a manifold M and the Riemannian metric given in local coordinates by the $g_{\mu\nu}$. If one is computing the expectation value of spectral observables, then instead of performing the integral over all pairs $(M, g_{\mu\nu})$ one can use the integration variable D. In fact, this suffices both to compute the observables and to define the spectral action. This does not imply that the map $(M, g_{\mu\nu}) \mapsto D$ is faithful or that it recovers the 4-manifold topology. The use of D as a "random Hermitian matrix" provides a simpler and more natural choice of an integration variable (neglecting constraints), than the $(M, g_{\mu\nu})$.

It is crucial to also include the matter fields in order to take into account not only gravity but the full coupled action

$$S = S_{EH} + S_{SM}$$

of gravity and matter. As we have seen in the previous sections, the key reason for dealing with noncommutative spaces is that a very simple modification of the geometry of space-time, obtained by taking the product with a finite noncommutative geometry F, provides the following results.

- The group of symmetries of gravity and matter, as in §§13.3 and 15.1.
- The gauge bosons, including the Higgs, as the inner fluctuations of the metric, as in §§15.2 and 15.4.
- The action $S = S_{EH} + S_{SM}$ as the spectral action

  $$\mathrm{Tr}(f(D/\Lambda)) + \frac{1}{2}\langle J\xi', D\xi\rangle$$

  as proved in Theorem 1.217.
- The free parameters of the Standard Model and of the gravitational terms all acquire geometric significance as shown in §17.12 above.

The parameters of the Standard Model that we obtain geometrically as summarized in §17.12 above are physical in both their signs and size but cannot have much significance until the renormalization group is brought into play. In order to achieve this one needs to test the spectral action under the renormalization group and the simplest way to proceed is to set up a functional integral using the spectral action in the exponent. Since the latter is spectral it admits a huge group of invariance, namely the group $\mathcal{U}$ of unitary operators on Hilbert space commuting with γ and J. We found in the geometric interpretation of the Standard Model that the KO-dimension of the product geometry $M \times F$ is equal to 2 modulo 8. Thus $J^2 = -1$ and the pair $i = \sqrt{-1}, J$ generate an action of the division algebra $\mathbb{H}$ of quaternions on the Hilbert space $\mathcal{H}$; the group $\mathcal{U}$ is thus the symplectic group. The first difficulty in the formulation of a functional integral is to specify precisely the integration variable, which roughly speaking should be the self-adjoint operator D corresponding to the most general noncommutative geometry of the correct dimension, since there is no reason to restrict a priori the amount of noncommutativity in the space-time coordinates. Since the inner fluctuations (§10.8) act on D in a linear way it is natural to take $\mathcal{D}[D]$ as the formal integration measure but one needs to write down the constraints fulfilled by D. It would be tempting, if one were just inspired by classical gravitation, to abstract the properties of a Dirac operator on an arbitrary 4-manifold. But we also learnt in the above discussion something quite new, which came from the matter sector, namely that we should deal with more subtle spaces, i.e. those having

- Metric dimension equal to 4.
- KO-dimension equal to $10 \sim 2$ (modulo 8).

This suggests that there might be a “raison d’être” for crossing by the space F just to get the right KO-dimension. After dealing with the main constraint i.e. the presence of a real orientation in §18.1 and the reconstruction Theorem in §18.2 we shall show indeed in §18.3 that the classification ([**49**], [**50**]) of irreducible finite noncommutative geometries of KO-dimension 6 quickly yields to the above finite geometry F. Finally we shall summarize the remaining difficulties for setting up a functional integral over all geometries, in §18.4.

18.1. Real orientation and volume form.

Let $(\mathcal{A}, \mathcal{H}, D, J, \gamma)$ be a real spectral triple and $\mathcal{B}$ the subalgebra (over $\mathbb{C}$) of $\mathcal{L}(\mathcal{H})$ generated by $\mathcal{A}$ and $J\mathcal{A}J^{-1}$. We let $\mathcal{A}_J = \{a \in \mathcal{A} \,|\, aJ = Ja\}$ as in §10.3 above. It is a subalgebra (over $\mathbb{R}$) of the center of $\mathcal{A}$ and commutes with $\mathcal{B}$.

Proposition 1.232. *For any integer n the following equality defines a map π from the Hochschild homology group $H_n(\mathcal{A}_J, \mathcal{B})$ to $\mathcal{L}(\mathcal{H})$,*

$$\begin{aligned} &\pi\left(\sum a_0(k) \otimes a_1(k) \otimes \cdots \otimes a_n(k)\right) \\ &= \sum a_0(k)[D, a_1(k)] \cdots [D, a_n(k)] \end{aligned} \tag{1.871}$$

Proof. By construction π is a linear map and is well defined on reduced chains since $[D, 1] = 0$. Let us show that its kernel contains the boundaries. These are given as sums of expressions of the form,

$$\begin{aligned} \sum_{j=0}^{n} \quad & (-1)^j a_0(k) \otimes \cdots \otimes a_j(k) a_{j+1}(k) \otimes \cdots \otimes a_{n+1}(k) \\ - \quad & (-1)^n a_{n+1}(k) a_0(k) \otimes \cdots \otimes \cdots \otimes a_n(k), \end{aligned}$$

where $a_0(k) \in \mathcal{B}$ and the $a_j(k) \in \mathcal{A}_J$ for $j > 0$. Their image under π is of the form

$$(-1)^n \left[a_0(k)[D, a_1(k)] \ldots [D, a_n(k)], a_{n+1}(k)\right].$$

Now by proposition 1.125 3) this is equal to 0 and we thus see that π factors through Hochschild boundaries. □

Definition 1.233. *Let $(\mathcal{A}, \mathcal{H}, D, J, \gamma)$ be a real spectral triple. A* real orientation *of dimension n is a Hochschild n-cycle $c \in H_n(\mathcal{A}_J, \mathcal{B})$ such that*

$$\text{even case}: \quad \pi(c) = \gamma\,, \quad \text{odd case}: \quad \pi(c) = 1\,.$$

Notice that the existence of a real orientation of dimension n implies that the Hochschild dimension of $\mathcal{A}_J$ is $\geq n$ since the class of c is non-zero by Proposition 1.232. Also notice that, unless D is bounded, one cannot have, in the even case, a real orientation of dimension 0. Indeed, if $\gamma = \pi(c)$, the equality

$$D = -\frac{1}{2}\,\gamma\,[D, \gamma] = -\frac{1}{2}\,c\,[D, c]$$

implies the boundedness of D, since we have $c \in \mathcal{B}$.

In the case of an ordinary compact oriented spin 4-manifold M, one obtains the cycle c as follows. One lets U_k be an open cover of M by domains of local coordinates $x_j(k)$ (extended smoothly outside U_k). Let $x(k)$ be a smooth partition of unity subordinated to the covering. Then, up to a normalization, one has

$$c = \sum_{k,\sigma} \epsilon(\sigma)\sqrt{g}\,x(k) \otimes x_{\sigma(1)}(k) \otimes \cdots \otimes x_{\sigma(4)}(k),$$

where the sum runs over all permutations σ of $\{1, \ldots, 4\}$ and $\epsilon(\sigma)$ is the signature of the permutation.

Let $(\mathcal{A}, \mathcal{H}, D, J, \gamma)$ be a real spectral triple of metric dimension 4, with a real orientation $c \in H_4(\mathcal{A}_J, \mathcal{B})$. We now explain the relation between c and the volume form i.e. the linear form which appears as the coefficient of Λ^4 in the spectral action:

$$a \in \mathcal{A} \mapsto ⨍ a\, D^{-4}\,. \tag{1.872}$$

We let ϕ be the cyclic 4-cocycle $\phi \in Z^4(\mathcal{B}, \mathcal{B})$ given by (cf. [**68**] IV.1.α)

$$\phi(a_0, \ldots, a_4) = \mathrm{Tr}'(\gamma\, a_0[F, a_1] \cdots [F, a_4])\,, \quad \forall a_j \in \mathcal{B}, \tag{1.873}$$

with $\mathrm{Tr}'(T) = \frac{1}{2}\mathrm{Tr}(F(FT + TF))$ and where the operator F is the sign of the operator D.

One then has the following homological interpretation of the linear form (1.872):

Theorem 1.234. *Let $(\mathcal{A}, \mathcal{H}, D, J, \gamma)$ be a real spectral triple of metric dimension 4, with a real orientation $c \in H_4(\mathcal{A}_J, \mathcal{B})$, then*

$$⨍ a\, D^{-4} = \langle ac, \phi\rangle\,, \quad \forall a \in \mathcal{A}\,. \tag{1.874}$$

where the right-hand side is the pairing between cycles and cocycles in the Hochschild cohomology.

Proof. By [**68**] IV.2.γ, Theorem 8, applied to the spectral triple

$$(\mathcal{B}, \mathcal{H}, D)$$

(with grading γ) one has for any Hochschild 4-cycle $v \in Z_4(\mathcal{B}, \mathcal{B})$, the equality

$$⨍ \gamma\pi(v) D^{-4} = \langle v, \phi\rangle\,, \tag{1.875}$$

where the map π is defined using the same formula as in (1.871). The commutation of $\mathcal{A}_J$ with $\mathcal{A}$ shows that the chain ac obtained by multiplication of the zeroth components of c by a is still a 4-cycle $ac \in Z_4(\mathcal{A}_J, \mathcal{B}) \subset Z_4(\mathcal{B}, \mathcal{B})$. Moreover the image of ac under the map π is $\pi(ac) = a\pi(c) = a\gamma = \gamma a$. Thus using (1.875), together with $\gamma^2 = 1$ one obtains the equality (1.874). $\square$

Remark 1.235. Both Proposition 1.232, Definition 1.233 and Theorem 1.234 extend to the following more general situation. Let $(\mathcal{A}, \mathcal{H}, D, \gamma)$ be an even spectral triple and J be an antilinear isometry fulfilling the properties of a real structure of Definition 1.124, but with the order one condition (1.473) weakened to the following

$$[[D, a], b^0] = 0 \qquad \forall\, a, b \in \mathcal{A}_J\,, \tag{1.876}$$

where $\mathcal{A}_J = \{a \in \mathcal{A} \,|\, aJ = Ja\}$ as in §10.3 above.

As discussed in [**275**] the orientability axiom does not hold for F in the usual formulation given in [**70**]. But by Corollary 1.176, it does however hold, in the stronger sense of Definition 1.233 and Remark 1.235, provided one uses the algebra $\mathcal{A}_{LR}$ instead of $\mathcal{A}_F$, and similarly $C^\infty(M, \mathcal{A}_{LR})$ instead of $C^\infty(M, \mathcal{A}_F)$.

18.2. The reconstruction of spin manifolds.

Besides having the correct metric dimension and fulfilling the regularity condition (1.490), and the orientability, the additional properties of the Dirac spectral triple (for $\mathcal{A} = C^\infty(M)$) which are used in the characterization given in [**70**] are the following.

(1) The space $\mathcal{H}^\infty = \cap_k \mathrm{Dom}(D^k)$ is a finite projective $\mathcal{A}$-module, endowed with an $\mathcal{A}$-valued inner product $\langle \xi, \eta \rangle_{\mathcal{A}}$ defined by

$$\langle a\xi, \eta \rangle = \int\!\!\!\!\!\!- a \, \langle \xi, \eta \rangle_{\mathcal{A}} \, |D|^{-n} \,, \quad \forall a \in \mathcal{A} \,, \xi \,, \eta \in \mathcal{H}$$

(2) The intersection form

$$K_*(\mathcal{A}) \times K_*(\mathcal{A}) \to \mathbb{Z}$$

obtained from the Fredholm index of D with coefficients in

$$K_*(\mathcal{A} \otimes \mathcal{A}^0)$$

is invertible.

The various solutions to the equations given by the conditions listed above are parameterized by the Riemannian metrics $g_{\mu\nu}$ and the spin structures on M. Each spin structure σ yields an affine space A_σ of solutions. The corresponding Dirac operator associated to the Levi–Civita connection of the metric $g_{\mu\nu}$ is the unique minimum in the affine space A_σ of the action functional which is given in dimension 4 by the residue

$$\int\!\!\!\!\!\!- D^{-2} \tag{1.877}$$

i.e. the Λ^2 coefficient of the spectral action of §11.1. This means that using the spectral action eliminates the extra degrees of freedom. This is similar to what happens in the form of Einstein gravity given by the Palatini action, where the basic dynamical variables are the tetrads (the vierbein) e^a_μ of (1.581) (see for instance [**7**] §3.2). It is nevertheless surely important to compute the full spectral action (including the terms independent of Λ) in the presence of the additional degrees of freedom.

The above characterization of the Dirac spectral triple (cf. [**132**] for the proof) on a spin manifold assumes that the algebra $\mathcal{A}$ is of the form $\mathcal{A} = C^\infty(M)$. We expressed the hope in [**70**] that, in the commutative case, the stated conditions for a real spectral triple single out not only the Dirac operators but also the smooth manifolds among general compact spaces. This latter hope is still unproved as formulated and in some sense it is not unexpected that the class of spaces thus obtained could contain a larger class of "pseudo-manifolds". These should nevertheless contribute to the functional integral. A precise reconstruction theorem for Riemannian spin manifolds under slightly different conditions was recently proved by Rennie and Varilly [**252**].

18.3. Irreducible finite geometries of KO-dimension 6.

We briefly explain a more conceptual path ([**49**], [**50**]) to the algebra $\mathcal{A}_F$ and its representation based on the classification of irreducible finite geometries of KO-dimension 6 modulo 8.
We start with a finite dimensional Hilbert space $\mathcal{H}$ endowed with an antiunitary operator J such that $J^2 = 1$. For any operator x in $\mathcal{H}$ we let, as in (1.472),

$$x^0 = Jx^*J^{-1}\,. \tag{1.878}$$

We look for involutive algebras $\mathcal{A}$ of operators in $\mathcal{H}$ such that (cf. (1.471)),

$$[x, y^0] = 0\,, \quad \forall x, y \in \mathcal{A}\,. \tag{1.879}$$

and that the following two conditions hold:

(1) The action of $\mathcal{A}$ has a separating vector, i.e. $\exists \xi \in \mathcal{H}$ such that $\mathcal{A}'\xi = \mathcal{H}$, where $\mathcal{A}'$ is the commutant of $\mathcal{A}$.
(2) The representation of $\mathcal{A}$ and J in $\mathcal{H}$ is irreducible.

The role of the first condition is to abstract a natural property of the action of an algebra of (smooth) functions on the sections of a vector bundle.
The meaning of the second condition is that one cannot find a non-trivial projection $e \in \mathcal{L}(\mathcal{H})$ which commutes with $\mathcal{A}$ and J.
We let $\mathcal{A}_{\mathbb{C}}$ be the complex linear space generated by $\mathcal{A}$ in the algebra $\mathcal{L}(\mathcal{H})$ of all operators in $\mathcal{H}$. We then have ([**49**]),

THEOREM 1.236. *Let $\mathcal{H}$ be a Hilbert space of dimension n. Then an irreducible solution $(\mathcal{A}, \mathcal{H}, J)$ exists iff one of the following two cases hold:*

- *$n = k^2$ is a square. Then the solution is given by $\mathcal{A}_{\mathbb{C}} = M_k(\mathbb{C})$ acting by left multiplication on itself and antilinear involution*

$$J(x) = x^*\,, \quad \forall x \in M_k(\mathbb{C})\,. \tag{1.880}$$

- *$n = 2k^2$ is twice a square. Then the solution is given by $\mathcal{A}_{\mathbb{C}} = M_k(\mathbb{C}) \oplus M_k(\mathbb{C})$ acting by left multiplication on itself and antilinear involution*

$$J(x, y) = (y^*, x^*)\,, \quad \forall x, y \in M_k(\mathbb{C})\,. \tag{1.881}$$

Moreover, by [**49**] Lemma 3.1, the commutation relation $J\gamma = -\gamma J$ coming from the hypothesis of KO-dimension 6 (cf. (1.470)) excludes the first case. Thus we concentrate on the second case. One can describe the solution in a more intrinsic manner as follows. We let V and W be k-dimensional complex Hilbert spaces. Then

$$\mathcal{A}_{\mathbb{C}} = \mathrm{End}_{\mathbb{C}}(W) \oplus \mathrm{End}_{\mathbb{C}}(V)\,. \tag{1.882}$$

We let $\mathcal{H}$ be the bimodule over $\mathcal{A}_{\mathbb{C}}$ given by

$$\mathcal{H} = \mathcal{E} \,\oplus\, \mathcal{E}^*\,, \quad J(\xi, \eta) = (\eta^*, \xi^*) \tag{1.883}$$

where,

$$\mathcal{E} = \mathrm{Hom}_{\mathbb{C}}(V, W)\,, \quad \mathcal{E}^* = \mathrm{Hom}_{\mathbb{C}}(W, V) \tag{1.884}$$

and the algebra acts on the left by composition:

$$(w, v)(g, h) = (w \circ g, v \circ h)\,, \quad \forall (w, v) \in \mathcal{A}_{\mathbb{C}}\,,\ (g, h) \in \mathcal{E} \,\oplus\, \mathcal{E}^*\,. \tag{1.885}$$

The various real forms i.e. the involutive algebras $\mathcal{A}$ with the above $\mathcal{A}_{\mathbb{C}}$ can then be described using additional antilinear isometries of V and W. The real forms of the algebra $M_k(\mathbb{C})$ in the above context are:

- $\mathcal{A} = M_k(\mathbb{C})$ (unitary case)
- $\mathcal{A} = M_k(\mathbb{R})$ (orthogonal case)
- $\mathcal{A} = M_a(\mathbb{H})$, for even $k = 2a$, (symplectic case)

For $\mathcal{A}_{\mathbb{C}} = M_k(\mathbb{C}) \oplus M_k(\mathbb{C})$ as in our case the list is more complicated but we make the *hypothesis* that both the grading and the real form come by assuming that the vector space W of (1.882) is a right vector space over $\mathbb{H}$ and is non-trivially $\mathbb{Z}/2$-graded. Since W is a non-trivially $\mathbb{Z}/2$-graded vector space over $\mathbb{H}$ its dimension must be at least 2 (and hence 4 when viewed as a complex vector space). We choose the simplest case i.e. W is a two-dimensional space over $\mathbb{H}$, and there is no ambiguity since all non-trivial $\mathbb{Z}/2$-gradings are equivalent. Our algebra is

$$\mathcal{A} = \mathrm{End}_{\mathbb{H}}(W) \oplus \mathrm{End}_{\mathbb{C}}(V) \sim M_2(\mathbb{H}) \oplus M_4(\mathbb{C})\,. \tag{1.886}$$

It follows from the grading of W that the algebra (1.886) is also $\mathbb{Z}/2$-graded, with non-trivial grading only on the $M_2(\mathbb{H})$-component. We still denote by γ the gradings of $\mathcal{E} = \mathrm{Hom}_{\mathbb{C}}(V, W)$ and $\mathcal{E}^* = \mathrm{Hom}_{\mathbb{C}}(W, V)$ given by composition with the grading of W. There exists up to equivalence a unique $\mathbb{Z}/2$-grading of $\mathcal{H}$ compatible with the graded representation of $\mathcal{A}$ and such that:

$$J\,\gamma = -\gamma\, J\,. \tag{1.887}$$

It is given by

$$\mathcal{H} = \mathcal{E} \,\ominus\, \mathcal{E}^*\,, \quad \gamma(\xi, \eta) = (\gamma\xi, -\gamma\eta)\,. \tag{1.888}$$

The center of our algebra $Z(\mathcal{A})$ is non-trivial and in that way the corresponding space is not connected. We look for "Dirac operators" D which connect non-trivially the two pieces (we call them "off-diagonal" as in §13.1) i.e. operators such that:

$$[D, Z(\mathcal{A})] \neq \{0\}. \tag{1.889}$$

The main requirement on such operators is the order one condition (1.473). We look for subalgebras $\mathcal{A}_F \subset \mathcal{A}^{\mathrm{ev}}$, the even part of $\mathcal{A}$, for which this order condition ($\forall\, a, b \in \mathcal{A}_F$) allows for operators which fulfill (1.889). We can now state the following analogue (cf. [**49**]) of Proposition 1.180 which recovers in a more conceptual manner the main "input" of §13.

THEOREM 1.237. *Up to an automorphism of $\mathcal{A}^{\mathrm{ev}}$, there exists a unique involutive subalgebra $\mathcal{A}_F \subset \mathcal{A}^{\mathrm{ev}}$ of maximal dimension admitting off-diagonal Dirac operators. It is given by*

$$\mathcal{A}_F = \{(\lambda \oplus q, \lambda \oplus m) \mid \lambda \in \mathbb{C}\,,\ q \in \mathbb{H}\,,\ m \in M_3(\mathbb{C})\} \subset \mathbb{H} \oplus \mathbb{H} \oplus M_4(\mathbb{C})\,,$$

using a field morphism $\mathbb{C} \to \mathbb{H}$. The involutive algebra $\mathcal{A}_F$ is isomorphic to $\mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C})$ and together with its representation in $(\mathcal{H}, J, \gamma)$ it gives (after tensoring by $\mathbb{C}^3$) the finite noncommutative geometry F of §13.

REMARK 1.238. Note that using the algebra $\mathcal{A} = \mathcal{A}_6$ of (1.886) (as a $\mathbb{Z}/2$-graded algebra) restores the Poincaré duality of §13.6. One finds that $K_0(\mathcal{A}) = \mathbb{Z}^2$ and that the pairing (1.692) is non-degenerate.

18.4. The functional integral and open questions.

We would like to formulate a functional integral over all "geometries" in the Euclidean form. In our framework a geometry is specified by a spectral triple i.e. a Hilbert space representation of the algebra $\mathcal{A}$ of coordinates and the inverse line element D. In first approximation we are interested in performing the functional integral as a sum over all spin 4-manifolds M and all metrics but since we need to take matter into account, we need to use not $C^\infty(M)$ as an algebra but rather the algebra $\mathcal{A} = C^\infty(M, \mathcal{A}_6)$, where $\mathcal{A}_6$ is given in (1.886).
Notice that the following set of data associated to a compact Riemannian spin 4-manifold (M, g) are in fact independent, up to isomorphism, of the choice of (M, g).

- The Hilbert space $\mathcal{H}_4$ of L^2-spinors.
- The $\mathbb{Z}/2$ grading operator $\gamma = \gamma_5$ on $\mathcal{H}_4$.
- The charge conjugation operator J_4.
- The decreasing filtration $\mathcal{H}_4^s \subset \mathcal{H}_4^{s'}$, $s > s'$ of L^2-spinors by Sobolev spaces.

The filtration is related to the notion of *rigged Hilbert space* and *Gelfand triple*, see [**138**].
Thus we can set the stage by tensoring the above by the finite geometry described in §18.3 before the reduction to $\mathcal{A}_F$. It is given by $(\mathcal{A}_6, \mathcal{H}_F, J_F, \gamma_F)$. At this point we have

- The Hilbert space $\mathcal{H} = \mathcal{H}_4 \otimes \mathcal{H}_F$ with its filtration by Sobolev spaces $\mathcal{H}^s$.
- The $\mathbb{Z}/2$ grading operator $\gamma = \gamma_5 \otimes \gamma_F$.
- The charge conjugation operator $J = J_4 \otimes J_F$.
- The small algebra $1 \otimes \mathcal{A}_6$ acting in $\mathcal{H}$.

The two ingredients we are missing are the algebra $\mathcal{A} \supset \mathcal{A}_6$ and the operator D. Now when $M = S^4$ is the 4-sphere, the algebra extension

$$C^\infty(M, M_4(\mathbb{C})) \supset M_4(\mathbb{C})$$

is generated by adding a projection $e = e^* = e^2$ fulfilling very simple conditions (cf. [**84**]). To see what happens one can view S^4 as the hypersurface in $\mathbb{R}^5 = \mathbb{R} \times \mathbb{H}$ given by the equation $t^2 - t + qq^* = 0$ and use the projection $e \in C^\infty(S^4, M_4(\mathbb{C}))$ given by

$$e = \begin{pmatrix} t & q \\ q^* & 1-t \end{pmatrix},$$

which together with the constant matrices generates $C^\infty(S^4, M_4(\mathbb{C}))$. Now in our case we are not dealing with $M_4(\mathbb{C})$ but with $\mathcal{A}_6 = M_2(\mathbb{H}) \oplus M_4(\mathbb{C})$. This means that if we simply adjoin to $\mathcal{A}_6$ a single projection e, fulfilling suitable commutativity conditions and in particular commuting with $Z(\mathcal{A}_6)$, we should expect to get two copies of S^4 and a correspondence between them, given by the bimodule $\mathcal{H}$. This might be enough to obtain all the algebras $\mathcal{A}$ of interest in the form $\mathcal{A} = \mathcal{A}_6(e)$. It remains to take care of the operator D, and write down the constraints it fulfills. The main relation is the *real orientability* of Definition 1.233 which gives a polynomial equation which relates D, c and γ, of the form

$$\langle c, D^{(4)} \rangle = \gamma, \tag{1.890}$$

where the pairing $\langle c, D^{(4)}\rangle$ is defined as in (1.871),

$$\begin{aligned} \langle c, D^{(4)}\rangle &= \sum a_0(k)[D,\, a_1(k)] \cdots [D,\, a_4(k)]\,, \\ c &= \sum a_0(k)\otimes a_1(k) \,\cdots \otimes\, a_4(k). \end{aligned} \tag{1.891}$$

The far distant goal is to set up a functional integral evaluating spectral observables $\mathcal{S}$ as

$$\langle \mathcal{S}\,\rangle = \mathcal{N} \int \mathcal{S}\; e^{-\mathrm{Tr}(f(D/\Lambda)) - \frac{1}{2}\langle J\psi,\, D\,\psi\rangle - \rho(e,D)}\, \mathcal{D}[\psi]\mathcal{D}[D]\mathcal{D}[e], \tag{1.892}$$

where $\gamma\psi = \psi$ is a Graßmann variable, and the term $\rho(e, D)$ implements the constraints which relate D and the algebra $\mathcal{A} = \mathcal{A}_6(e)$.

Remark 1.239. The question of "quantizing" the Hochschild cycle c of (1.890) by writing it as the Chern character of an idempotent e motivated the results of [**84**], [**78**], [**79**], [**80**] on the structure of "noncommutative spheres" characterized by the following condition. One requires that

$$\exists\; e = e^2 = e^* \in M_q(\mathcal{A})\,, \quad c = \mathrm{Ch}_2(e)$$

while the lower components $\mathrm{Ch}_j(e)$ $(j = 0, 1)$ of the Chern character $\mathrm{Ch}(e)$ of the projection e are all required to vanish. The algebra $\mathcal{A}$ is then generated by the matrix components of e and it is quite remarkable that, besides the usual sphere, this problem singles out a large class of noncommutative manifolds.

19. Dimensional regularization and noncommutative geometry

The aim of this section is to construct tools for the computation of chiral anomalies [**15**] in the general framework of noncommutative geometry. We show that a careful investigation of the meaning of the Breitenlohner-Maison prescription for the compatibility of the chiral symmetry with the procedure of Dimensional Regularization leads us to an interpretation of DimReg, as far as one loop fermionic graphs are concerned, not just as a formal procedure, but as an actual *geometry in complexified dimension $D - z$*, based on noncommutative spaces and spectral triples. We then obtain a description of the chiral anomalies in terms of spectral triples and the local index formula of §10.5.

Our starting point is the physical prescription of how to use DimReg in the presence of spinor fields.

19.1. Chiral anomalies.

The cubic anomaly is central to the problem of renormalizability of a theory like the Standard Model, which involves a gauge theory. The problem is that, in such a theory, one needs to ensure that the renormalized Lagrangian is still invariant under the same group of local gauge symmetries that leave the bare Lagrangian invariant. Thus, it is necessary, in order to have renormalizability and unitarity, that gauge invariance is preserved at each order in the renormalized perturbation series.

Dimensional Regularization (DimReg) is considered the best regularization method, when dealing with the problem of gauge invariance, because it does not change the formal expression of the Ward identities, while analytically continuing the dimension $D = 4$ to complex points $D - z = d \in \mathbb{C}$.

The main advantage of this regularization procedure is that it is *canonical*, in the sense that it respects all the symmetries such as space-time (Lorentz invariance) or gauge symmetries. However, this property fails in the presence of spinor fields, as in that case the chiral symmetry is not preserved. The main problem is how one should treat the γ_5 in DimReg. In fact, in dimension $D = 4$, the symbol γ_5 indicates the product of the gamma matrices,

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3, \tag{1.893}$$

where the γ^μ satisfy the Clifford relations

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\, I, \quad \text{with} \quad \mathrm{Tr}(I) = 4, \tag{1.894}$$

and γ_5 anticommutes with them,

$$\{\gamma_5, \gamma^\mu\} = 0. \tag{1.895}$$

When one complexifies the dimension around a critical dimension D, the naive prescription which formally sets γ_5 to still anticommute with symbols γ^μ while keeping the cyclicity of the trace is not consistent and produces contradictions ([**62**], §13.2). There is a more sophisticated prescription of 't Hooft-Veltman [**169**] and Breitenlohner-Maison [**34**] on how to make sense of the γ_5 in DimReg.

In general the Ward identities determine relations between the Green's functions arising from symmetries of the Lagrangian. These influence the renormalizability of theories with nontrivial symmetries, by ensuring cancellations of divergences between different sectors of the theory. There is a complicated interplay of symmetry and renormalization, whereby the Ward identities can be affected by higher order corrections, which can generate anomalous terms. In gauge theories, the problem arises primarily from fermions coupling to gauge fields. This manifests itself in the Adler–Bell–Jackiw (ABJ) anomaly for electroweak interactions.
The ABJ anomaly arises in the QED Ward identity involving axial-vector currents, in the renormalization of 1-loop diagrams (triangle fermion loops) [**1**], [**15**].
Let j_μ, $j_{\mu,5}$, and P denote, respectively, the vector, axial, and pseudoscalar currents

$$\begin{aligned} j_\mu(x) = & \; \overline{\psi}(x)\gamma_\mu\psi(x) \\ j_{\mu,5}(x) = & \; \overline{\psi}(x)\gamma_\mu\gamma_5\psi(x) \\ P(x) = & \; \overline{\psi}(x)\gamma_5\psi(x). \end{aligned} \tag{1.896}$$

Let $T_{\mu\nu\lambda}$ and $T_{\mu\nu}$ denote the 3-point functions, with $q = k_1 + k_2$,

$$\begin{aligned} T_{\mu\nu\lambda}(k_1, k_2) = & \; i \textstyle\int d^4x_1\, d^4x_2\, \langle 0|T(j_\mu(x_1)j_\nu(x_2)j_{\lambda,5}(0))|0\rangle\, e^{ik_1\cdot x_1 + ik_2\cdot x_2} \\ T_{\mu\nu}(k_1, k_2) = & \; i \textstyle\int d^4x_1\, d^4x_2\, \langle 0|T(j_\mu(x_1)j_\nu(x_2)P(0))|0\rangle\, e^{ik_1\cdot x_1 + ik_2\cdot x_2}. \end{aligned}$$

Then, at tree level, the vector and axial-vector Ward identities are given by

$$k_1^\mu T_{\mu\nu\lambda} = k_2^\nu T_{\mu\nu\lambda} = 0 \quad \text{and} \quad q^\lambda T_{\mu\nu\lambda} = 2mT_{\mu\nu}, \tag{1.897}$$

where m is the mass of the fermion field ψ. These are derived from the equations of motion

$$\partial^\mu j_\mu = 0 \quad \text{and} \quad \partial^\mu j_{\mu,5} = 2imP. \tag{1.898}$$

The contributions to $T_{\mu\nu\lambda}$ and $T_{\mu\nu}$ from the triangle diagrams with internal fermion lines are given by

$$\begin{aligned} T_{\mu\nu\lambda} = & \tfrac{-i}{(2\pi)^4} \int d^4p \ \Big(\mathrm{Tr} \Big(\tfrac{i}{\not p - m} \gamma_\lambda \gamma_5 \tfrac{i}{(\not p - \not q) - m} \gamma_\nu \tfrac{i}{(\not p - \not k_1) - m} \gamma_\mu \Big) \\ & + \mathrm{Tr} \Big(\tfrac{i}{\not p - m} \gamma_\lambda \gamma_5 \tfrac{i}{(\not p - \not q) - m} \gamma_\mu \tfrac{i}{(\not p - \not k_2) - m} \gamma_\nu \Big) \Big) \end{aligned} \tag{1.899}$$

$$\begin{aligned} T_{\mu\nu} = & \tfrac{-i}{(2\pi)^4} \int d^4p \ \Big(\mathrm{Tr} \Big(\tfrac{i}{\not p - m} \gamma_5 \tfrac{i}{(\not p - \not q) - m} \gamma_\nu \tfrac{i}{(\not p - \not k_1) - m} \gamma_\mu \Big) \\ & + \mathrm{Tr} \Big(\tfrac{i}{\not p - m} \gamma_5 \tfrac{i}{(\not p - \not q) - m} \gamma_\mu \tfrac{i}{(\not p - \not k_2) - m} \gamma_\nu \Big) \Big). \end{aligned} \tag{1.900}$$

This generates an anomaly $q^\lambda T_{\mu\nu\lambda} = 2mT_{\mu\nu} + \Delta_{\mu\nu}$ for the axial-vector Ward identity, where the anomaly can be computed as

$$\Delta_{\mu\nu} = \frac{-1}{4\pi^2} \epsilon_{\mu\nu\sigma\rho} k_1^\sigma k_2^\rho, \tag{1.901}$$

by imposing the condition that the vector Ward identity is preserved. The anomaly (1.901) corresponds to the modification of the axial-vector current equation in (1.898) to

$$\partial^\mu j_{\mu,5} = 2imP - \frac{1}{(4\pi)^2} \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma},$$

with F the electromagnetic field.
One way to derive this anomaly is by change of variables by translations in divergent integrals, while another way, through DimReg, is more suitable for generalization to non-abelian cases and the full Standard Model. The ABJ anomaly can also be formulated in terms of path integrals as the fact that the measure is not invariant under the action of γ_5. Higher order terms do not contribute any further corrections.

In the DimReg scheme, the ABJ anomaly is closely related to the problem of defining γ_5 for complex dimensions. In fact, while for diagrams with open fermion lines it is possible to use a formal definition of γ_5 in arbitrary complex dimension, which satisfies anticommutation relations with all the γ_μ (cf. e.g. [**62**]), when one considers cases like the triangle diagram, with closed fermion loops, one runs into the further problem of making this formal definition of γ_5 compatible with the trace.
Notice that the ABJ anomaly is independent of the fermion mass. However, in the DimReg computation, when there are overlapping divergences, the leading term depends on the mass (cf. [**134**]).
For non-abelian theories, one can give an analogous derivation of the ABJ anomaly via DimReg. In this case the three masses associated to the three fermion lines may be different (the matrices in the group representation have off-diagonal terms) and one obtains (see e.g. [**57**]) an expression of the form

$$D^{abc} = \frac{1}{2} \mathrm{Tr}(\{T^a, T^b\} T^c), \tag{1.902}$$

so that, for the triangle diagram, the vanishing of the ABJ anomaly is given by the condition $D^{abc} = 0$.
In this non-abelian case more diagrams (up to five-points functions) contribute anomalous terms, but the triangle diagrams still remains the fundamental one, as the vanishing of that anomaly implies the absence of all other anomalies as well.

Massless QCD is chiral invariant at the classical level, but this invariance is broken, already at the classical level, by the presence of quark mass terms, in which case the anomaly in the axial symmetry appears in the form ([**62**], (13.1.5))

$$\partial^\mu j_{\mu,5} = -2i\bar{\psi}\gamma_5 M\psi + \text{anomaly}.$$

In the case of electroweak interactions where the vector boson associated to the weak interaction acquires mass through the Higgs mechanism, renormalizability is obtained by showing that the theory is obtained via a formal transformation of the fields from a renormalizable gauge theory (the latter can have non-physical particles that may spoil unitarity, cf. [**168**]). For such a formal transformation to be possible, one needs to be able to renormalize the gauge theory in a gauge invariant way. This can be spoiled by the presence of the ABJ anomaly, as shown in [**152**].

The question of constructing gauge theories that are free of anomalies, both for electroweak and strong interactions, was considered in [**141**], where conditions are given on which gauge groups and representations will satisfy the vanishing of (1.902). In the standard electroweak theory with $\mathrm{SU}(2)\times\mathrm{U}(1)$, if Y denotes the hypercharge, the condition that (1.902) vanishes for the fermion representation implies relations between the hypercharges. Since the group SU(2) is anomaly free (cf. [**141**]), at least one of the T in (1.902) has to be the U(1) hypercharge Y. For the case of one hypercharge Y, the vanishing of the anomaly for the triangle diagram yields a linear relation between the hypercharges $\sum_{leptons} Y + \sum_{quark} Y = 0$. The anomaly vanishes when two of the T's in (1.902) are equal to Y, while in the case of three hypercharges one obtains a cubic relation. More precisely, for $Y(\nu_{eL}) = Y(e_L) = -1$, $Y(e_R) = -2$, $Y(u_L) = Y(d_L) = 1/3$, $Y(u_R) = 4/3$, and $Y(d_R) = -2/3$, one has

$$\sum_{leptons} Y = -1\times 2 - 2 = -4 \quad \text{and} \quad \sum_{quark} Y = 3(\frac{1}{3}\times 2 + \frac{4}{3} - \frac{2}{3}) = 4,$$

where the factor 3 counts the color degrees of freedom, and

$$\sum_{\text{leptons}} (Q^2T_3 - QT_3^2) = -\frac{1}{2} + \frac{1}{4} = -\frac{1}{4}$$

$$\sum_{\text{quarks}} (Q^2T_3 - QT_3^2) = \frac{2}{3} - \frac{1}{6} - \frac{1}{2} + \frac{1}{4} = \frac{1}{4},$$

where $Y = 2(Q - T_3)$.

In the case of gauge theories, the Ward identities (or Slavnov–Taylor identities) are related to the symmetries by Becchi–Rouet–Stora (BRS) gauge transformations.

An interesting aspect of the equations imposed by the vanishing of the cubic anomalies is that it imposes constraints on the hypercharges and on the Weyl representations. In fact, it is shown in [**140**] and [**232**] that requiring the theory to be free of anomalies suffices to fix these data uniquely. This could be an interesting point if one wants to use the vanishing of the anomalies as a *constraint on the geometry*. In fact, it is shown in [**140**] that requiring the vanishing of all the known anomalies in four dimensions, namely the perturbative triangular chiral gauge anomalies and the mixed perturbative chiral gauge–gravitational anomaly discussed above, as well as the non-perturbative global SU(2) chiral gauge anomaly of [**300**], fixes uniquely both the massless fermion Weyl representation and the hypercharges (prior to spontaneous symmetry breaking of the electroweak subgroup).

19.2. The spaces X_z.

We look for a spectral triple whose inverse line element $D = D_z$ fulfills the following basic equation,

$$\mathrm{Tr}(e^{-\lambda D^2}) = \pi^{z/2}\lambda^{-z/2}, \quad \forall \lambda \in \mathbb{R}_+^*\,, \tag{1.903}$$

corresponding, using (1.124), to

$$\mathrm{Tr}(e^{-\lambda D^2}) = \int e^{-\lambda\, k^2}\, d^z k \quad \forall \lambda \in \mathbb{R}_+^*\,. \tag{1.904}$$

Let us first explain how to perform the construction in the slightly more general framework of type II spectral triples [**16**], [**39**], [**40**]. The only condition which is altered (with respect to Definition 1.120) in the definition of such a triple is the compactness of the resolvent of D which only holds relative to a type II subfactor N of the commutant of the algebra generated by $\mathcal{A}, D$ (and γ, J in the even or real cases). One then replaces everywhere the trace Tr by the type II trace Tr_N relative to the commutant of N. Since the analogues of the analysis of compact operators and of their ideals is well developed in the type II situation one obtains a theory ([**16**], [**39**], [**40**]) which can be used in a very similar way as the theory of spectral triples.

To construct D_z we let Z be a self-adjoint operator affiliated to a type II_∞ factor N and with spectral measure given by

$$\mathrm{Tr}_N(1_E(Z)) = \frac{1}{2}\int_E dy \tag{1.905}$$

for any interval $E \subset \mathbb{R}$, where 1_E is the characteristic function of E. We shall give in Chapter 3 Remark 3.36 a concrete realization, of arithmetic nature, for the pair (N, Z).

We then let, for any complex number $z \in \mathbb{C}$ with $\Re(z) > 0$,

$$D_z = \rho(z)\, F\, |Z|^{1/z} \tag{1.906}$$

where F is the sign of the operator Z, $|Z|$ its absolute value, and the complex power $|Z|^{1/z}$ is taken by the usual functional calculus. We take the normalization constant $\rho(z)$ to be

$$\rho(z) = \pi^{-\frac{1}{2}}\, (\Gamma(\frac{z}{2} + 1))^{\frac{1}{z}}. \tag{1.907}$$

PROPOSITION 1.240.
- *The operator D_z of* (1.906)*, with the normalization factor* (1.907)*, satisfies* (1.903)*, provided $z \in \mathbb{R}_+$.*
- *For $\Re(z) > 0$ and after imposing an infrared cutoff on the trace Tr_N the ζ-function of D_z has a single (simple) pole at $s = z$ and is absolutely convergent in the half space $\Re(s/z) > 1$.*

PROOF. 1) We assume $z \in \mathbb{R}_+^*$. One has

$$\mathrm{Tr}_N(e^{-\lambda D_z^2}) = \frac{1}{2}\int e^{-\lambda\, \rho^2\, |y|^{2/z}}\, dy.$$

Thus, by setting $\rho^2\, |y|^{2/z} = u$, one gets $dy = \rho^{-z}\, \frac{z}{2}\, u^{\frac{z}{2}-1}\, du$ and

$$\mathrm{Tr}_N(e^{-\lambda D_z^2}) = \rho^{-z}\, \frac{z}{2}\int_0^\infty e^{-\lambda\, u}\, u^{\frac{z}{2}-1}\, du = \rho^{-z}\, \Gamma(\frac{z}{2} + 1)\, \lambda^{-z/2},$$

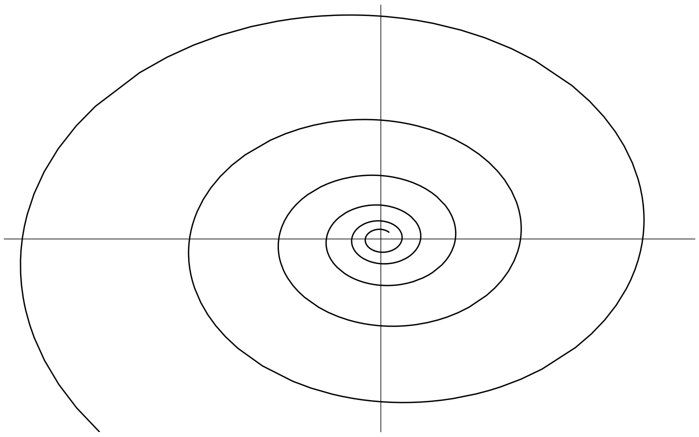


FIGURE 39. The spectrum of D_z^2 for $\Im(z) \neq 0$.

so that one obtains as required

$$\mathrm{Tr}_N(e^{-\lambda D_z^2}) = \pi^{z/2}\,\lambda^{-z/2}, \quad \forall \lambda \in \mathbb{R}_+^*. \tag{1.908}$$

Notice that $\rho(z)$ is just a normalization factor and it is well behaved in the limit $z \to 0$, since we have

$$\frac{1}{z}\,\log\Gamma(\frac{z}{2}+1) = -\frac{\gamma}{2} + \frac{\pi^2}{48}\,z + \cdots + (-1)^n \frac{\zeta(n)}{n\,2^n}\,z^{n-1} + \cdots \tag{1.909}$$

In particular, we have $\rho(0) = e^{-\gamma/2}$. (Here γ denotes the Euler constant, not to be confused with the γ used everywhere else for the grading!)
2) The above computation of $\mathrm{Tr}_N(e^{-\lambda D_z^2})$ makes perfect sense when $z \in \mathbb{R}_+^*$, but it requires more care when the imaginary part of z is non-zero. In fact, in that case the spectrum of D_z^2 is a spiral in $\mathbb{C}$ (cf. Figure 39), so that $e^{-\lambda D_z^2}$ no longer belongs to the domain of the trace Tr_N. One can define the trace by means of a suitable regularization, but it is simpler to directly deal with the ζ-function

$$\mathrm{Tr}_N((D_z^2)^{-s/2}) \tag{1.910}$$

instead of the theta function

$$\mathrm{Tr}_N(e^{-\lambda D_z^2}). \tag{1.911}$$

We need to impose an infrared cutoff, that is, we perform the integral

$$\mathrm{Tr}_N((D_z^2)^{-s/2}) = \frac{1}{2}\int (\rho^2\,|y|^{2/z})^{-s/2}\,dy \tag{1.912}$$

in the region outside $|y| < 1$. This gives, using Tr'_N to indicate the cutoff,

$$\mathrm{Tr}'_N\left((D_z^2)^{-s/2}\right) = \rho^{-s}\int_1^\infty u^{-s/z}\,du = \rho^{-s}\frac{z}{s-z} \tag{1.913}$$

which, as a function of s, has a single (simple) pole at $s = z$ and is absolutely convergent in the half space $\Re(s/z) > 1$. $\square$

The algebra $\mathcal{A}'$ for the spectral triple defining X_z will play no role below except for the unit element $1 \in \mathcal{A}'$. One could include in $\mathcal{A}'$ any operator a such that $[D_z, a]$ is bounded and both a and $[D_z, a]$ are smooth for the "geodesic flow"

$$T \mapsto e^{it|D_z|}\, T\, e^{-it|D_z|} \tag{1.914}$$

and it is an interesting open question to give an upper bound for the Minkowski dimension of the spectrum of such operators (the natural guess for the bound is $1/\Re(1/z) = x + y^2/x$ for $z = x + iy$).

The dimension spectrum of X_z is reduced to the single point z as shown by (1.913).

In order to justify the use of the infrared cutoff in (1.912), one can make the following observation. Let us replace the operator $|Z|$ by an operator of the form $f(|Z|)$ where $f \in C^\infty(\mathbb{R}^+)$ has the following properties:

- $f(x) = x$ for $x \geq 1/2$
- f is strictly increasing
- $f(0) > 0$.

Then the replacement $|Z| \to f(|Z|)$ alters (1.908) by the error term

$$R(\lambda, z) = \int_0^{\frac{1}{2}} (e^{-\lambda\, \rho^2\, |y|^{2/z}} - e^{-\lambda\, \rho^2\, f(|y|)^{2/z}})\, dy. \tag{1.915}$$

This is an analytic function of λ. In the region $\Re(z) > 0$, $|z| < 1$ and $|\lambda| < 2^{\Re(2/z)}$, it satisfies the estimate

$$|R(\lambda, z)| < C\, |\lambda|\, 2^{\Re(-2/z)}, \tag{1.916}$$

for some finite constant $C > 0$. This does not affect the pole parts in z for all computations involving DimReg, as we check below. Thus, replacing $D_z = \rho(z) F |Z|^{1/z}$ by $\tilde{D}_z = \rho(z) F f(|Z|)^{1/z}$ the ζ-function $\mathrm{Trace}_N((\tilde{D}_z^2)^{-s/2})$ is now well defined and it differs from (1.913) by an entire function of s which does not alter the dimension spectrum.

19.3. Chiral gauge transformations.

When the basic algebra $\mathcal{A}$ is a $\mathbb{Z}/2$-graded algebra, the discussion of §10.8 adapts by replacing everywhere the commutators $[D, a]$ with *graded* commutators

$$[D, a]_- := Da - (-1)^{\deg(a)} aD\,, \tag{1.917}$$

where the subscript $_-$ reminds one of the use of graded commutators. Also, with $\theta \in \mathrm{Aut}(\mathcal{A})$ the $\mathbb{Z}/2$-grading automorphism, the gauge transformations (1.508) are now of the form $\gamma_u(A) = \theta(u) du^* + \theta(u) A u^*$.

This applies in particular when one replaces the original algebra $\mathcal{A}$ with the algebra $\tilde{\mathcal{A}}$ generated (in the even case) by $\mathcal{A}$ and γ. As an algebra, $\tilde{\mathcal{A}}$ is isomorphic to the direct sum of two copies of $\mathcal{A}$ by

$$a + b\gamma \mapsto (a + b,\, a - b).$$

We endow it with the $\mathbb{Z}/2$-grading $\theta \in \mathrm{Aut}(\tilde{\mathcal{A}})$ such that $\theta(\gamma) = -\gamma$, while θ is the identity on $\mathcal{A}$. The unitary group $\tilde{\mathcal{U}}$ of $\tilde{\mathcal{A}}$ now contains γ and it is the group of *chiral* gauge transformations. However, the anticommutation of γ with D shows that this procedure does not generate really new gauge potentials, but just multiplies the existing ones by an arbitrary power of γ.

The situation becomes much more interesting when one keeps the algebra $\tilde{\mathcal{A}}$, but one replaces the original spectral triple by its product (in the sense of noncommutative geometry) with the space X_z.
Thus, we start with an even spectral triple $(\mathcal{A}, \mathcal{H}, D)$ with $\mathbb{Z}/2$-grading γ and form its product with the (type II_∞) spectral triple defining the space X_z. We therefore let $D' = D_z$, and we adopt the following shorthand notation (compatible with the physics notation of [**62**])

$$\bar{D} = D \otimes 1\,, \quad \hat{D} = \gamma \otimes D'\,, \quad D'' = \bar{D} + \hat{D}\,. \tag{1.918}$$

The fact that the spectral triple defining the space X_z is of type II_∞ introduces some technical subtleties that can be ignored at first reading (cf. [**16**], [**39**], [**40**] for the detailed general theory).
We let the algebra $\mathcal{A}$ act on $\mathcal{H} \otimes \mathcal{H}'$ by

$$a \mapsto a \otimes 1 \tag{1.919}$$

Thus, one obtains a type II_∞ spectral triple for the algebra $\mathcal{A}$. We then extend the algebra to the graded algebra $\tilde{\mathcal{A}}$ generated by $\mathcal{A}$ and γ, with the action on $\mathcal{H} \otimes \mathcal{H}'$ still given by (1.919).
The odd element $\gamma \in \tilde{\mathcal{A}}$ no longer anticommutes (i.e. graded-commutes) with D'', since $\gamma \otimes 1$ commutes with $\hat{D} = \gamma \otimes D'$. Thus, one gets a new non-trivial gauge potential of the form

$$B = [D'', \gamma]_- = 2\gamma \hat{D}\,. \tag{1.920}$$

This gauge potential accounts for the lack of invariance of D'' under the "chiral" gauge transformation $\gamma \in \tilde{\mathcal{U}}$ and it corresponds, in the language of QFT, to the divergence $\partial_\mu j_5^\mu$ of the axial current j_5^μ.
One needs special care in discussing this point, since it involves Euclidean Fermi fields ξ and η, which one needs to treat as independent integration variables in the Euclidean functional integral (see [**61**], §5.2 of "The use of instantons").
Thus, the fermionic part of the Lagrangian in the action functional is of the form

$$\mathcal{L}_{\text{fermions}} = \langle \eta\,,\, D'' \xi \rangle \tag{1.921}$$

When extending the theory to $\mathbb{Z}/2$-graded algebras, one defines the gauge transformations α_u on the fermions, for $u \in \tilde{\mathcal{U}}$, using the $\mathbb{Z}/2$-grading θ as

$$\alpha_u(\xi) = u\,\xi\,, \quad \alpha_u(\eta) = \theta(u)\,\eta\,. \tag{1.922}$$

It follows that the lack of invariance of the action (1.921) is still accounted for by the gauge potential given by the graded commutator with u. In particular if γ (which is odd) would anticommute with D'', then the action (1.921) would be invariant under the corresponding chiral transformation. This is not the case and the variation at order one of $\mathcal{L}_{\text{fermions}}$ under the chiral gauge transformation $u = e^{i\omega\gamma}$, for $\omega = \omega^* \in \mathcal{A}$, is given by

$$\delta\mathcal{L}_{\text{fermions}} = \langle \eta,\, (i\,[D,\omega]\,\gamma + i\,\omega B)\,\xi \rangle \tag{1.923}$$

with $B = 2\gamma\hat{D}$ as in (1.920).
We use the physics terminology *evanescent* to qualify the gauge potentials of the form $E = \omega B$. The origin of the terminology is clear since they vanish when the extra dimension z is set to 0.

19.4. Finiteness of anomalous graphs and relation with residues.

We fix throughout a regular even spectral triple $(\mathcal{A}, \mathcal{H}, D)$. We assume for simplicity that D is invertible and refer to (1.493) for the ϵ-trick allowing to remove this hypothesis. We denote by $OP(\mathcal{A}, \mathcal{H}, D)$ the algebra generated by $\mathcal{A}$, D and γ. The following lemma will suffice to prove the finiteness of the anomalous graphs.

LEMMA 1.241. *Suppose given a regular spectral triple $(\mathcal{A}, \mathcal{H}, D)$. Let P be an element of $OP(\mathcal{A}, \mathcal{H}, D)$. For $n > k > 0$ and in the limit $z \to 0$, one has*

$$\operatorname{Tr}(\hat{D}^{2k}\,(P \otimes 1)\,D''^{-2n}) = -\frac{1}{2}\,B(k, n-k) ⨍ P\,D^{-2(n-k)},$$

$$B(p,q) = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)}.$$

PROOF. For $z \neq 0$ one has

$$\operatorname{Tr}(\hat{D}^{2k}\,(P \otimes 1)\,D''^{-2n}) = \frac{1}{\Gamma(n)} \int_0^\infty \operatorname{Tr}(\hat{D}^{2k}\,(P \otimes 1)\,e^{-t\bar{D}^2}\,e^{-t\hat{D}^2})\,t^{n-1}\,dt$$

We also have

$$\operatorname{Tr}_N(D_z^{2k}\,e^{-tD_z^2}) = \frac{z\,(z+2)\cdots(z+2k-2)}{2^k}\,\pi^{z/2}\,t^{-z/2-k}, \tag{1.924}$$

for all $t \in \mathbb{R}_+^*$, while

$$\int_0^\infty e^{-tD^2}\,t^{n-1-z/2-k}\,dt = \Gamma(n - z/2 - k)\,|D|^{z-2(n-k)}. \tag{1.925}$$

One has $\hat{D}^{2k}\,(P \otimes 1) = P \otimes D_z^{2k}$ thus using the factorization of the trace and (1.924) and (1.925), we get

$$\operatorname{Tr}(\hat{D}^{2k}\,(P \otimes 1)\,D''^{-2n}) =$$

$$\pi^{z/2}\,\frac{\prod_0^{k-1}(z/2+j)}{\Gamma(n)}\,\Gamma(n - z/2 - k)\,\operatorname{Tr}(P\,|D|^{z-2(n-k)}),$$

which gives the required result since when $z \to 0$ one has,

$$\pi^{z/2}\,\frac{\prod_0^{k-1}(z/2+j)}{\Gamma(n)}\,\Gamma(n - z/2 - k) \sim B(k, n-k)\,\frac{z}{2},$$

with the minus sign coming from the distinction between $|D|^z$ and $|D|^{-z}$. ∎

As a corollary of Lemma 1.241 one gets the following expression for the residue at the simple pole $z = 0$ in the DimReg expression of one-loop graphs with only fermionic internal lines; we work as above with a regular spectral triple with simple dimension spectrum.

PROPOSITION 1.242. *Let $A \in \Omega_D^1$ be a gauge potential and $n > 0$. Then the expression*

$$\operatorname{Tr}(((A \otimes 1)\,D''^{-1})^n) \tag{1.926}$$

has at most a simple pole at $z = 0$ with residue given by

$$\operatorname{Res}_{z=0} \operatorname{Tr}(((A \otimes 1)\,D''^{-1})^n) = - ⨍ (A\,D^{-1})^n \tag{1.927}$$

PROOF. One has

$$D''^{-1} = D'' D''^{-2} = (\bar{D} + \hat{D}) D''^{-2} \tag{1.928}$$

and

$$D''^2 = \bar{D}^2 + \hat{D}^2. \tag{1.929}$$

Using Lemma 1.136 (2) to move the terms D''^{-2} to the right we can express (1.926) as a sum of terms of the form

$$\mathrm{Tr}((A \otimes 1)(\bar{D} + \hat{D})(A^{(k_1)} \otimes 1)(\bar{D} + \hat{D})) \cdots (A^{(k_n)} \otimes 1)(\bar{D} + \hat{D}) D''^{-2k}) \tag{1.930}$$

where we use the notation $A^{(m)} = \nabla^m(A)$ and $k = n + \sum k_j$. Only finitely many of these terms have a divergent trace at $z = 0$. By construction $\hat{D}$ anticommutes with the A's and commutes with all other terms. Thus one can expand (1.930) in terms of the form

$$\mathrm{Tr}\,(\hat{D}^{2q}\, P\, D''^{-2m})\,.$$

By Lemma 1.241 any of these terms with $q > 0$ is regular at $z = 0$. Thus the only contribution to the pole comes from the case $q = 0$. This means that one only takes the terms $\bar{D}$ in the expansion of (1.930) and one is thus dealing with

$$\mathrm{Tr}((P \otimes 1)\, D''^{-2k})\,, \quad P = A\, D\, \cdots A^{(k_j)}\, D \cdots A^{(k_n)}\, D \tag{1.931}$$

As in the proof of Lemma 1.241 one has

$$\mathrm{Tr}\,((P \otimes 1)\, D''^{-2k}) = \frac{1}{\Gamma(k)} \int_0^\infty \mathrm{Tr}((P \otimes 1)\, e^{-t\bar{D}^2}\, e^{-t\hat{D}^2})\, t^{k-1}\, dt$$

Using the factorization of the trace and (1.903) while

$$\int_0^\infty e^{-tD^2}\, t^{k-1-z/2}\, dt = \Gamma(k - z/2)\, |D|^{z-2k}$$

we get

$$\mathrm{Tr}\,((P \otimes 1)\, D''^{-2k}) = \pi^{z/2}\, \frac{\Gamma(k - z/2)}{\Gamma(k)}\, \mathrm{Tr}(P\, |D|^{z-2k})\,.$$

Thus the residue at $z = 0$ is given by

$$-\!\!\!\!\!\!\int P\, |D|^{-2k}$$

The computation of $-\!\!\!\!\!\!\int (A\, D^{-1})^n$ gives exactly the same result. Indeed, one again writes

$$D^{-1} = D\, D^{-2}$$

and uses Lemma 1.136 (2) to move the terms D^{-2} to the right, which makes it possible to express $-\!\!\!\!\!\!\int (A\, D^{-1})^n$ as a sum of terms of the form

$$-\!\!\!\!\!\!\int A\, D\, \cdots A^{(k_j)}\, D \cdots A^{(k_n)}\, D\, D^{-2k}$$

which match exactly with the above terms. $\square$

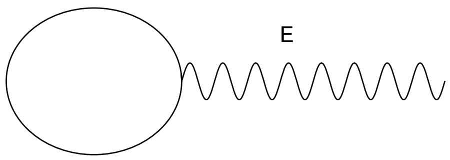


FIGURE 40. The tadpole.

19.5. The simplest anomalous graphs.

We now analyze the graphs with only fermionic internal lines which involve linearly the evanescent gauge potential B, starting from the simplest, which is the tadpole, and continuing to more involved cases.
The data we start with are, as above, an even spectral triple $(\mathcal{A}, \mathcal{H}, D)$ with simple dimension spectrum, i.e. we use exactly the same setup as in the local index formula of [**92**] that we recalled above in §10.5.

19.5.1. *The Tadpole.*

We start with the simplest graph. This has one loop with a single external leg to which is assigned an evanescent gauge potential $E = \gamma a\hat{D}$ (see Figure 40). The analytic expression for the tadpole is of the form

$$\mathrm{Tr}(E\, D''^{-1}), \tag{1.932}$$

where D''^{-1} plays the role of the fermionic propagator. We assume $⨍\gamma\, a = 0$ for all $a \in \mathcal{A}$ and obtain the following result.

LEMMA 1.243. *Let φ_0 be the linear form* (1.495), *which is the zeroth order component of the local index cocycle, as in Theorem 1.134. In the limit $z \to 0$, one obtains*

$$\mathrm{Tr}\,(E\, D''^{-1}) = -\,\varphi_0(a), \quad \forall a \in \mathcal{A}\,. \tag{1.933}$$

PROOF. One has

$$D''^{-2} = \int_0^\infty e^{-t\bar{D}^2}\, e^{-t\hat{D}^2}\, dt. \tag{1.934}$$

Thus, for fixed z, one gets using (1.928) and (1.929),

$$\mathrm{Tr}(E\, D''^{-1}) = \mathrm{Tr}(\gamma\, a\hat{D}\,(\bar{D} + \hat{D})\, D''^{-2}) = \mathrm{Tr}(\gamma\, a\hat{D}^2\, D''^{-2}). \tag{1.935}$$

In fact, the terms with an odd number of $\hat{D}$'s have zero trace since one can find an involution in $\mathcal{H}'$ which anticommutes with D_z. Using (1.934) this gives

$$\mathrm{Tr}(E\, D''^{-1}) = \int_0^\infty \mathrm{Tr}(\gamma\, a\hat{D}^2\, e^{-t\bar{D}^2}\, e^{-t\hat{D}^2})\, dt \tag{1.936}$$

and the trace factorizes to give

$$\mathrm{Tr}(\gamma\, a\hat{D}^2\, e^{-tD^2}\, e^{-t\hat{D}^2}) = \mathrm{Tr}(\gamma\, a\, e^{-tD^2})\mathrm{Tr}(D_z^2\, e^{-tD_z^2}). \tag{1.937}$$

By the basic equality (1.908) one has

$$\mathrm{Tr}_N(e^{-tD_z^2}) = \pi^{z/2} t^{-z/2}, \quad \forall t \in \mathbb{R}_+^*\,. \tag{1.938}$$

After differentiating in t this gives

$$\mathrm{Tr}_N(D_z^2\, e^{-tD_z^2}) = \frac{z}{2}\, \pi^{z/2}\, t^{-z/2-1} \quad \forall t \in \mathbb{R}_+^*\,. \tag{1.939}$$

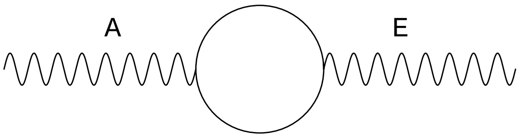


FIGURE 41. The self-energy graph.

Thus, (1.936) and (1.937) give

$$\mathrm{Tr}(E\,D''^{-1}) = \frac{z}{2}\,\pi^{z/2}\int_0^\infty \mathrm{Tr}(\gamma\,a\,e^{-tD^2}\,)\,t^{-z/2-1}\,dt \tag{1.940}$$

Moreover, we have

$$\int_0^\infty e^{-tD^2}\,t^{-z/2-1}\,dt = \Gamma(-z/2)\,|D|^z, \tag{1.941}$$

while the limit for $z \to 0$ of $\mathrm{Tr}(\gamma\,a\,|D|^z)$ is $\varphi_0(a)$. Thus, we get the required equality since, for $z \to 0$, we have

$$\frac{z}{2}\,\pi^{z/2}\Gamma(-z/2) \to -1.$$

□

19.5.2. *The self-energy graph.*

The next term linear in E now involves a usual gauge potential

$$A = \sum a_i\,[D'',\,b_i]$$

(see Figure 41). Its analytic expression is given by

$$\mathrm{Tr}(E\,D''^{-1} A\,D''^{-1}). \tag{1.942}$$

One has the following result.

LEMMA 1.244. *In the limit $z \to 0$, one obtains*

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}) = \sum_0^\infty (-1)^{n+1}\frac{1}{2n+2}\,⨍\,\gamma\,a\,\nabla^n(B)\,D^{-2n-2} \tag{1.943}$$

where $B = dA + A'$ with

$$dA = \sum\,[D,a_i]\,[D,\,b_i]\,, \qquad A' = \sum\,a_i\,\nabla(b_i)\,.$$

and only finitely many terms in the infinite sum are non-zero.

PROOF. Using (1.928), one obtains

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}) = \mathrm{Tr}(\gamma\,a\hat{D}\,(\bar{D} + \,\hat{D})\,D''^{-2}\,A\,(\bar{D} + \,\hat{D})\,D''^{-2})\,.$$

Since the terms with an odd number of $\hat{D}$ do not contribute, this gives

$$\mathrm{Tr}(\gamma\,a\hat{D}^2\,D''^{-2}\,A\,\bar{D}\,D''^{-2}) + \mathrm{Tr}(\gamma\,a\hat{D}\,\bar{D}\,D''^{-2}\,A\,\hat{D}\,D''^{-2})$$

In the second term the second appearance of $\hat{D}$ can be put in front provided one takes care of signs, since $\hat{D}$ commutes with D''^{-2} and anticommutes with both $A = \sum a_i\,[D'',\,b_i] = \sum a_i\,[\bar{D},\,b_i]$ and $\bar{D}$.

Thus, we can write the above as

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}) = \mathrm{Tr}(\gamma\,a\hat{D}^2\,D''^{-2}\,(A\,\bar{D}\,+\bar{D}\,A)\,D''^{-2})\,.$$

One has

$$(A\,\bar{D}\,+\bar{D}\,A) = B\otimes 1\,,\qquad B = dA + A'\,, \tag{1.944}$$

where

$$dA = \sum\,[D,a_i]\,[D,\,b_i]\,,\quad A' = \sum\,a_i\,\nabla(b_i)\,. \tag{1.945}$$

Thus we can use Lemma 1.136 to move the first D''^{-2} across and get (only finitely many terms will contribute)

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}) = \sum_0^\infty(-1)^n\,\mathrm{Tr}(\gamma\,a\hat{D}^2\,(\nabla^n(B)\otimes 1)\,D''^{-2n-4})\,.$$

Since $\hat{D}^2$ commutes with $a\in\mathcal{A}$ we then conclude using Lemma 1.241.

□

Notice that, since the regularized trace is a trace outside the dimension spectrum, the order of the terms $E\,D''^{-1}$ and $A\,D''^{-1}$ is irrelevant.
It is worthwhile to double check this in this example by redoing the calculation with the other order. One has

$$\begin{aligned}
\mathrm{Tr}(A\,D''^{-1}\,E\,D''^{-1}) &= \mathrm{Tr}(A\,(\bar{D}+\,\hat{D})\,D''^{-2}\,\gamma\,a\hat{D}\,(\bar{D}+\,\hat{D})\,D''^{-2})\\
&= \mathrm{Tr}(A\,\hat{D}\,D''^{-2}\,\gamma\,a\hat{D}\,\bar{D}\,D''^{-2}) + \mathrm{Tr}(A\,\bar{D}\,D''^{-2}\,\gamma\,a\hat{D}^2\,D''^{-2})\\
&= -\,\mathrm{Tr}(\gamma\,\hat{D}^2\,A\,D''^{-2}\,a\,\bar{D}\,D''^{-2}) + \mathrm{Tr}(\gamma\,\hat{D}^2\,A\,D''^{-2}\,\bar{D}\,a\,D''^{-2})\\
&= \sum_0^\infty(-1)^n\,\mathrm{Tr}(\gamma\,\hat{D}^2\,A\,(\nabla^n([D,a])\otimes 1)\,D''^{-2n-4})\\
&= \sum_0^\infty(-1)^{n+1}\frac{1}{2n+2}\,⨍\,\gamma\,A\,\nabla^n([D,a])\,D^{-2n-2}\,.
\end{aligned}$$

We thus get

$$\mathrm{Tr}(A\,D''^{-1}\,E\,D''^{-1}) = \sum_0^\infty(-1)^{n+1}\frac{1}{2n+2}\,⨍\,\gamma\,A\,\nabla^n([D,a])\,D^{-2n-2}\,. \tag{1.946}$$

It might seem at first sight that one passes from (1.943) to (1.946) by a simple integration by parts, i.e. using

$$⨍\,\gamma\,\nabla^n(B)\,C = (-1)^n\,⨍\,\gamma\,B\,\nabla^n(C) \tag{1.947}$$

but it is a bit more subtle. One uses the equality

$$D^{-2n-2}\,a = \sum_0^\infty\,(-1)^k\,c(n,k)\,\nabla^k(a)\,D^{-2n-2k-2} \tag{1.948}$$

with

$$c(n,0) = 1\,,\quad c(n,k) = \frac{\prod_1^k\,(n+j)}{k!}\,, \tag{1.949}$$

to write

$$\int\!\!\!\!\!- \gamma\, a\, \nabla^n(B)\, D^{-2n-2} = \sum_0^\infty (-1)^k\, c(n,k) \int\!\!\!\!\!- \gamma\, \nabla^n(B)\, \nabla^k(a)\, D^{-2n-2k-2}$$

which after integration by parts yields

$$\int\!\!\!\!\!- \gamma\, a\, \nabla^n(B)\, D^{-2n-2} = \sum_0^\infty (-1)^{n+k}\, c(n,k) \int\!\!\!\!\!- \gamma\, B\, \nabla^{(n+k)}(a)\, D^{-2n-2k-2}$$

The desired equality between (1.943) and (1.946) then follows using

$$\sum_0^m (-1)^k \frac{c(m-k,k)}{m-k+1} = \int_0^1 (t-1)^m\, dt = (-1)^m \frac{1}{m+1}$$

and the other integration by parts

$$\int\!\!\!\!\!- \gamma\, (DA + AD)\, C = \int\!\!\!\!\!- \gamma\, A\, [D, C]\,. \tag{1.950}$$

To conclude we thus have the two equivalent expressions (1.943) and (1.946) for the self-energy graph.
We can express both in a more compact notation using the derivation (cf. [**92**]),

$$\Theta(T) = \sum_1^\infty \frac{(-1)^{n+1}}{n}\, \nabla^n(T)\, D^{-2n} \tag{1.951}$$

which generates the one-parameter group Ad D^{2z}. Indeed one has

$$D^2\, T\, D^{-2} = T + \nabla(T)\, D^{-2} = (1+\epsilon)(T)$$

and by construction

$$\Theta = \log(1+\epsilon)\,, \quad e^\Theta = 1+\epsilon\,. \tag{1.952}$$

We also need the basic “Planck” function

$$\pi(z) = \frac{z}{e^z - 1} \tag{1.953}$$

which is familiar in index theory in the formulas giving the characteristic classes.
We can then rewrite the above formulas as follows.

Proposition 1.245. *One has, in the limit $z \to 0$,*

$$\begin{aligned} \mathrm{Tr}(E\, D''^{-1}\, A\, D''^{-1}) &= -\tfrac{1}{2} \int\!\!\!\!\!- \gamma\, a\, \pi(\Theta)(B)\, D^{-2} \\ &= -\tfrac{1}{2} \int\!\!\!\!\!- \gamma\, A\, \pi(\Theta)([D,a])\, D^{-2} \end{aligned}$$

Proof. From the above it is enough to check that

$$\pi(\Theta) = \sum_0^\infty \frac{(-1)^n}{n+1}\, \epsilon^n \tag{1.954}$$

which follows from (1.952) and (1.953). $\square$

It also follows from the above discussion that for X and Y in the algebra generated by $\mathcal{A}$, γ and D one has the formula of integration by parts

$$\int\!\!\!\!\!- X\,\pi(\Theta)(Y)\,D^{-2} = \int\!\!\!\!\!- Y\,\pi(\Theta)(X)\,D^{-2}\,. \tag{1.955}$$

One can understand this formula as an instance of the KMS condition fulfilled by the functional

$$\varphi_1(X) = \int\!\!\!\!\!- X\,D^{-2} \tag{1.956}$$

with respect to the one-parameter group of automorphisms $\sigma_t = e^{i\,t\,\Theta}$. Indeed one has, as required by the KMS_1 condition

$$\varphi_1(X\,\sigma_i(Y)) = \varphi_1(Y\,X)$$

and (1.955) follows from the simple equality

$$\pi(-z) = e^z\,\pi(z)\,,$$

combined with the invariance of φ_1 under σ_t which implies that for any formal series f,

$$\varphi_1(f(\Theta)(X)\,Y) = \varphi_1(X\,f(-\Theta)(Y))\,. \tag{1.957}$$

Recall from §10.4 that a Hochschild cochain φ of dimension n on an algebra $\mathcal{A}$, defines (cf. [**68**]) a functional on the universal n-forms $\Omega^n(\mathcal{A})$ by (1.482):

$$\int_\varphi a_0\,da_1\cdots\,da_n = \varphi(a_0,\,a_1,\ldots,a_n)$$

and when φ is a Hochschild cocycle one has the trace property (1.483), i.e.

$$\int_\varphi a\,\omega = \int_\varphi \omega\,a, \quad \forall a\in\mathcal{A}$$

Moreover the boundary operator B_0 of (1.485) is defined in such a way that one has the analogue (1.486) of Stokes formula

$$\int_\varphi d\,\omega = \int_{B_0\varphi} \omega$$

Proposition 1.245 makes it possible to express the self-energy graph in terms of the following two-cochain

$$\psi_2(a_0,a_1,a_2) = \frac{1}{4}\int\!\!\!\!\!- \gamma\,a_0\,[D,a_1]\,\pi(\Theta)([D,a_2])\,D^{-2} \tag{1.958}$$

so that

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}) = -2\int_{\psi_2} A\,da \tag{1.959}$$

PROPOSITION 1.246. (1) *The cochain $B_0\psi_2$ is cyclic :*

$$A\,B_0\psi_2 = 2\,B_0\psi_2.$$

(2) $B\,\psi_2 + b\,\varphi_0 = 0$

(3) *If* $b\,\varphi_0 = 0$ *one has*

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}) = -2\int_{\psi_2}(dA)a \tag{1.960}$$

PROOF. (1) One has

$$B_0\psi_2(a_0,a_1) = \frac{1}{4}\fint \gamma\,[D,a_0]\,\pi(\Theta)([D,a_1])\,D^{-2}$$

which is antisymmetric by (1.955) (γ anticommutes with $[D,a]$).
(2) Since D anticommutes with γ and $D[D,a]+[D,a]D = \nabla(a)$ one gets

$$B\,\psi_2(a_0,a_1) = -\frac{1}{2}\fint \gamma\,a_0\,\pi(\Theta)(\nabla(a_1))\,D^{-2} = -\frac{1}{2}\fint \gamma\,a_0\,\Theta(a_1)$$

The coboundary $b\varphi_0$ is given by

$$b\varphi_0(a,b) = \mathrm{Lim}_{s\to 0}\,\mathrm{Tr}(\gamma\,a(b\,|D|^{-s} - |D|^{-s}\,b)) =$$

$$\mathrm{Lim}_{s\to 0}\,\mathrm{Tr}(\gamma\,a\,(1-e^{-\frac{s}{2}\Theta})(b)\,|D|^{-s}) = \frac{1}{2}\fint \gamma\,a\,\Theta(b)$$

(3) In that case one has, by (1) and (2), $B_0(\psi_2) = 0$ and

$$\fint_{\psi_2}(dA)a - \fint_{\psi_2}A\,da = \fint_{\psi_2}d(Aa) = 0$$

by (1.486) thus the result follows from (1.959). ∎

19.6. Anomalous graphs in dimension 2 and the local index cocycle.

We deal with the two-dimensional case and explain how the result of Proposition 1.246 simplifies in this case. The main result of this section is Theorem 1.250 which shows that the sum of the anomalous graphs is given by the pairing with the local index cocycle.
In the two-dimensional case, the top component of the local index cocycle is given by the following Hochschild 2-cocycle:

$$\varphi_2(a_0,\,a_1,a_2) = \frac{1}{4}\fint \gamma\,a_0\,[D,a_1]\,[D,a_2]\,D^{-2}, \quad \forall a_j\in\mathcal{A}. \tag{1.961}$$

Since one is in dimension 2 all the terms with $n>0$ in (1.954) vanish and one gets $\varphi_2 = \psi_2$.

In case the coboundary of the tadpole vanishes, i.e. assuming that $b\varphi_0 = 0$ (1.933), one gets by Theorem 1.134 that φ_2 is a cyclic cocycle and hence $\int_{\varphi_2}$ is a closed graded trace (cf. §10.4). We thus get,

COROLLARY 1.247. *In dimension* ≤ 2, *and if* $b\varphi_0 = 0$, φ_2 *is a cyclic cocycle and*

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}) = -2\int_{\varphi_2}a\,dA$$

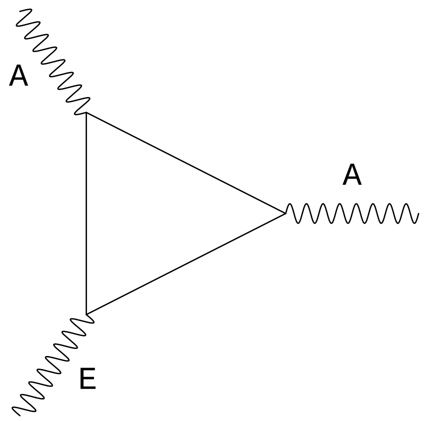


FIGURE 42. The triangle graph.

Next, we consider the original ABJ triangle graph. The computation below can prepare the ground for the 4-dimensional case and we thus do it in the required generality.
The analytic expression for the triangle graph of Figure 42 is given by

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}\,A\,D''^{-1}). \tag{1.962}$$

This can be polarized in the following manner:

$$\mathrm{Tr}(E\,D''^{-1}\,A_1\,D''^{-1}\,A_2\,D''^{-1}), \tag{1.963}$$

where one can now assume that the A_j are monomials of the form

$$A_j = a_j\,[D, b_j]\,.$$

We first prove the following lemma, independent of the dimension.

LEMMA 1.248. *Let $P_j \in OP(\mathcal{A}, \mathcal{H}, D)$, for $k = 1$ and $k = 2$ one has in the limit $z \to 0$,*

$$\mathrm{Tr}(\hat{D}^{2k}\,P_0\,D''^{-2}\,P_1\,D''^{-2}\,P_2\,D''^{-2}) =$$
$$\sum c(a,b,k) ⨍ P_0 \nabla^a(P_1)\,\nabla^b(P_2)\,D^{-2(a+b+3-k)}$$

where

$$c(a,b,k) = \frac{1}{2}\,(-1)^{a+b+1}\,\frac{(k-1)!(a+b+2-k)!}{(a+b+2)\,b!\,(a+1)!}$$

PROOF. One has, using Lemma 1.136, for $P_j \in OP(\mathcal{A}, \mathcal{H}, D)$,

$$D''^{-2}\,P_1\,D''^{-2}\,P_2 \sim \sum d(a,b)\,\nabla^a(P_1)\,\nabla^b(P_2)\,D''^{-2(a+b+2)}$$

where the coefficients are given by

$$d(a,b) = (-1)^{a+b} \sum_{0\leq c\leq b} \frac{(a+c)!}{a!\,c!} = (-1)^{a+b}\,\frac{(a+b+1)!}{b!\,(a+1)!}$$

Thus using Lemma 1.241 for $n = a+b+3$ and the equality

$$\frac{1}{2}\,\frac{(a+b+1)!}{b!\,(a+1)!}\,\frac{(k-1)!(a+b+2-k)!}{(a+b+2)!} = \frac{1}{2}\,\frac{(k-1)!(a+b+2-k)!}{(a+b+2)\,b!\,(a+1)!}$$

one gets the required formula. □

Notice that one has

$$D^{-2}\,P_1\,D^{-2}\,P_2 \sim \sum d(a,b)\,\nabla^a(P_1)\,\nabla^b(P_2)\,D^{-2(a+b+2)}$$

but this does not allow us to undo the above reordering under the residue since different powers of D'' occur in applying Lemma 1.241.

PROPOSITION 1.249. *In the two-dimensional case one has in the limit $z \to 0$ (with $E = \gamma a \hat{D}$)*

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}\,A\,D''^{-1}) = 2\int_{\varphi_2} a\,A^2\,.$$

PROOF. Using (1.928), one gets

$$\mathrm{Tr}(E\,D''^{-1}\,A\,D''^{-1}\,A\,D''^{-1}) =$$
$$\mathrm{Tr}(\gamma\,a\hat{D}\,(\bar{D}+\,\hat{D})\,D''^{-2}\,A\,(\bar{D}+\,\hat{D})\,D''^{-2}\,A\,(\bar{D}+\,\hat{D})\,D''^{-2})\,.$$

Since the terms with an odd number of $\hat{D}$ do not contribute we get terms with 4 occurrences of $\hat{D}$, i.e.

$$\begin{aligned} T_4 =&\ \mathrm{Tr}(\gamma\,a\hat{D}^2\,D''^{-2}\,A\,\hat{D}\,D''^{-2}\,A\,\hat{D}\,D''^{-2}) \\ =&\ -\mathrm{Tr}(\gamma\,a\hat{D}^4\,D''^{-2}\,A\,D''^{-2}\,A\,D''^{-2}) \end{aligned}$$

(where the minus sign comes from the anticommutation of $\hat{D}$ with A) and terms with two occurrences of $\hat{D}$ which give the following terms T_j, $j \le 3$:

$$\begin{aligned} T_1 =&\ \mathrm{Tr}(\gamma\,a\hat{D}^2\,D''^{-2}\,A\,\bar{D}\,D''^{-2}\,A\,\bar{D}\,D''^{-2}) \\ T_2 =&\ \mathrm{Tr}(\gamma\,a\hat{D}\,\bar{D}\,D''^{-2}\,A\,\hat{D}\,D''^{-2}\,A\,\bar{D}\,D''^{-2}) \\ =&\ \mathrm{Tr}(\gamma\,a\hat{D}^2\,\bar{D}\,D''^{-2}\,A\,D''^{-2}\,A\,\bar{D}\,D''^{-2}) \\ T_3 =&\ \mathrm{Tr}(\gamma\,a\hat{D}\,\bar{D}\,D''^{-2}\,A\,\bar{D}\,D''^{-2}\,A\,\hat{D}\,D''^{-2}) \\ =&\ \mathrm{Tr}(\gamma\,a\hat{D}^2\,\bar{D}\,D''^{-2}\,A\,\bar{D}\,D''^{-2}\,A\,D''^{-2}) \end{aligned}$$

Let us first compute T_4 and T_2. By Lemma 1.248, we get

$$T_4 = \frac{1}{4}\,⨍\,\gamma\,a\,A^2\,D^{-2}$$

since all the terms with non-zero powers of ∇ give zero since we are in dimension 2. Next, by Lemma 1.248, we get

$$T_2 = -\frac{1}{4}\,⨍\,\gamma\,a\,D\,A^2\,D^{-3}$$

We can replace $a\,D \mapsto D\,a$ since the commutator $[D,a]$ is bounded. Thus since the sign changes when we permute D and γ we get, using the trace property of the residue,

$$T_2 = \frac{1}{4}\,⨍\,\gamma\,a\,A^2\,D^{-2}$$

We thus get

$$T_2 + T_4 = 2\int_{\varphi_2} a\,A^2$$

Next one has by Lemma 1.248,

$$T_1 = -\frac{1}{4} ⨍ \gamma\, a\, A\, DAD^{-3}$$

Since one is in dimension 2 one can permute the factor D^{-2} which gives

$$T_1 = -\frac{1}{4} ⨍ \gamma\, a\, A\, D^{-1} A\, D^{-1}$$

If we let

$$\psi(a_0, a_1) = ⨍ \gamma\, a_0\, [D, a_1]\, D^{-1}$$

we get

$$\begin{aligned} b\psi(a_0, a_1, a_2) = & \; -⨍ \gamma\, a_0\, [D, a_1]\, [a_2, D^{-1}] \\ = & \; -⨍ \gamma\, a_0\, [D, a_1]\, D^{-1}\, [D, a_2]\, D^{-1} \end{aligned}$$

Moreover one has

$$\int_{b\psi} a\, A^2 = -⨍ \gamma\, a\, A\, D^{-1} A\, D^{-1}$$

as one checks replacing the first A by $a_1[D, b_1]$ and the second by $a_2[D, b_2]$, and showing that

$$\begin{aligned} ⨍ \gamma\, a\, a_1[D, b_1]\, D^{-1}\, a_2[D, b_2]\, D^{-1} &= ⨍ \gamma\, a\, a_1[D, b_1] a_2\, D^{-1}\, [D, b_2]\, D^{-1} \\ &= -\int_{b\psi} a\, a_1\, db_1\, a_2\, db_2 \end{aligned}$$

We have thus shown that

$$T_1 = \frac{1}{4} \int_{b\psi} a\, A^2$$

Since $b\psi$ is a coboundary this ensures that T_1 is not a significant term but in fact we now show that it is exactly canceled by T_3; indeed one has

$$\begin{aligned} T_3 = -\frac{1}{4} ⨍ \gamma\, a\, D\, A\, D\, A\, D^{-4} &= -\frac{1}{4} ⨍ \gamma\, D\, a\, A\, D\, A\, D^{-4} \\ &= \frac{1}{4} ⨍ \gamma\, a\, A\, D\, A\, D^{-3} = -T_1 \end{aligned}$$

where we used the boundedness of $[D, a]$ to permute D and a and the minus sign comes from permuting D with γ. ☐

We can now summarize the relation with the index cocycle in the two-dimensional case as follows,

THEOREM 1.250. *Assume that the dimension is ≤ 2 and that the coboundary of the tadpole vanishes, $b\varphi_0 = 0$; then*

1. *φ_2 is a cyclic two-cocycle.*
2. *For any $a \in \mathcal{A}$ and any gauge potential A one has*

$$\mathrm{Tr}(E\, D''^{-1}\, A\, D''^{-1}\, A\, D''^{-1}) - \mathrm{Tr}(E\, D''^{-1}\, A\, D''^{-1}) = 2 \int_{\varphi_2} a\, (dA + A^2)$$

In other words the alternating sum of the anomalous graphs has a simple interpretation in terms of the pairing of the curvature $F = dA + A^2$ and of the local index cocycle.

Proof. This follows from Proposition 1.249 and Corollary 1.247. $\square$

CHAPTER 2

The Riemann zeta function and noncommutative geometry

1. Introduction

This chapter describes, following the results of [**71**], a spectral realization of the zeros of the Riemann zeta function and an interpretation of Weil's explicit formulae of number theory as a trace formula. The chapter also includes an application of the same techniques to the Archimedean local factors of L-functions of varieties defined over number fields, following [**74**].

We begin the chapter by recalling some basic facts about the Riemann zeta function. In §2 we recall the fundamental relation between primes and zeros of the Riemann zeta function expressed by Riemann in terms of an explicit formula for the prime counting function. We also recall Riemann's estimate for the counting of the nontrivial zeros of the zeta function.

In §3 we describe a classical Hamiltonian system (the "scaling Hamiltonian") and a corresponding quantum mechanical system that recover Riemann's estimate of the counting of zeros of zeta as a counting of modes of the physical system. More precisely, in §3.1 we show a striking similarity between the behavior of the oscillatory part in Riemann's formula for the counting function and the corresponding oscillatory part in the semi-classical formula for the number of eigenvalues of the Hamiltonian obtained by quantizing a classical Hamiltonian system. The two formulae are easily matched up to a sign, which suggests the important point of interpreting the spectrum of our scaling system as an "absorption" rather than an "emission" spectrum. This distinction will play a crucial role both in the results described in this chapter as well as in the formulation given in Chapter 4, where the spectral realization and the trace formula we discuss will live naturally on a space that is a cokernel (or more precisely, as will become evident in Chapter 4) a *motive* in an abelian category of noncommutative spaces. We proceed in §3.2 to introduce a classical Hamiltonian system associated to the group of scaling transformations of the real line. We show that the symplectic volume with infrared and ultraviolet cutoff gives the average part of the Riemann counting function. In §3.3 we describe the quantization of this classical system. A delicate point in obtaining a counting of the modes (the energy levels) of the resulting quantum system lies in the implementation of both an ultraviolet and an infrared cutoff, for the reason that one cannot impose a cutoff on both a function and its Fourier transform, or to say it more precisely, the two projectors associated to the two types of cutoff do not commute, hence one cannot implement them simultaneously by intersecting their ranges. The problem can be solved using a technique that was developed in electrical engineering and laser technology to deal precisely with similar sorts of mathematical problems.

This is based on the existence of a differential operator commuting with both cutoff projections, whose eigenfunctions, the *prolate spheroidal wave functions*, can be used to approximate both cutoffs by restricting to a subspace spanned by a number of them, depending on the energy range allowed by the cutoff. We then consider the spectral projections of the scaling action, so that the problem of the counting of quantum states becomes the computation of the trace of the product of these spectral projections N_E with the projection Q_Λ on the span of the spheroidal wave functions implementing the cutoff. The trace can be computed (as a special case of the more general trace formula proved in §7.2 of this same chapter). The resulting trace $\mathrm{Tr}(Q_\Lambda N_E)$ is expressed in a distributional form in terms of a principal value.

The definition and properties of such principal values are described in detail in the following section, §4, following [**71**]. We begin by giving the formal computation of the distributional trace of operators of the form

$$\vartheta_{\mathrm{a}}(h) = \int_{K^*} h(\lambda)\vartheta_{\mathrm{a}}(\lambda)\, d^*\lambda$$

with h a test function and ϑ_{a} the scaling action of $\mathbb{G}_m(K)$ on L^2 functions on $\mathbb{G}_a(K)$ by

$$(\vartheta_{\mathrm{a}}(\lambda)\,\xi)(q) = \xi(\lambda^{-1}q),$$

which covers the case of the operators we used in §3.3 to describe the spectral projections of the scaling action. We give in §4.1 a quick introduction to *modulated groups* and the corresponding Haar measures, while in §4.2 we use these notions to give the general form of principal values as distributions extending from $\mathbb{G}_m(K)$ to $\mathbb{G}_a(K)$ the integration with respect to the normalized Haar measure. In particular we give a description in terms of a variant of the "minimal subtraction" method of regularization described in Chapter 1 (see Definition 1.10).

In §5 we come to the actual counting of the modes of the quantum scaling Hamiltonian. We show in Theorem 2.18 that the trace $\mathrm{Tr}(Q_\Lambda N_E)$ gives the right expression for the counting that matches the leading term in Riemann's formula. The proof of this theorem is given in §5.2 using the noncommutative geometry method known as *quantized calculus*, which we briefly review in §5.1.

We then come in §6 to the mathematical implementation of the "absorption spectrum" idea. We introduce the map $\mathfrak{E}$ considered in [**71**], which acts as

$$\mathfrak{E}(f)(\lambda) = \lambda^{1/2}\sum_{n\in\mathbb{Z}} f(n\lambda)$$

for $\lambda \in \mathbb{R}_+^*$. We show in §6.1 that, after approximating spheroidal wave functions by Hermite–Weber functions, the Fourier transform $\widehat{\mathfrak{E}(f)}$ is the product of an entire function by the Riemann ξ function.

In §7 we make an important extension of the quantum mechanical system described by the Hamiltonian associated to scaling transformations on the real line to an *adelic* version where one considers all the places of $\mathbb{Q}$ and not only the Archimedean place, thus replacing $\mathbb{R}$ by the adèles $\mathbb{A}_\mathbb{Q} = \mathbb{A}_{\mathbb{Q},f} \times \mathbb{R}$ and $\mathbb{R}^*$ acting by scaling by $\mathrm{GL}_1(\mathbb{A}_\mathbb{Q})$ acting by multiplication. Thus, the Hilbert space of the quantum statistical mechanical system becomes $L^2(\mathbb{A}_\mathbb{Q}/\mathbb{Q}^*)$ with the action of the idèle class group $C_\mathbb{Q} = \mathbb{A}_\mathbb{Q}^*/\mathbb{Q}^*$. The quotient space $X_\mathbb{Q} = \mathbb{A}_\mathbb{Q}/\mathbb{Q}^*$ is called the *adèle class space* and, as we see in more detail in Chapter 3 and Chapter 4, it should be regarded as a noncommutative space. While in this chapter we mostly treat $X_{\mathbb{Q}}$ as an ordinary space, we explain in §7.1 the reason why it is more natural to regard it as a noncommutative space, due to the ergodic nature of the action of $C_{\mathbb{Q}}$. We explain this in the case of the restriction to a set S of finitely many places (semi-local case $X_{\mathbb{Q},S} = \mathbb{A}_{\mathbb{Q},S}/\mathbb{Q}_S^*$). Illustrative examples of two or three places are discussed explicitly in §§7.1.1 and 7.1.2. We explain in §7.2 how to define on rapidly decaying functions on $\mathbb{A}_{\mathbb{Q},S}$ an inner product that descends to the coinvariants of the action of $\mathbb{Q}_S^*$ and we use it to define the Hilbert space $L^2(X_{\mathbb{Q},S})$ with the corresponding action of $C_{\mathbb{Q},S}$. We also extend the proof of the trace formula given via quantized calculus from the case of the single Archimedean place to the case of finitely many places, obtaining in this way the semi-local trace formula of [**71**]. A main feature of the trace formula is its *additivity*: even though the adèle class space is essentially a product over the set $\Sigma_{\mathbb{K}}$ of places of the global field $\mathbb{K}$ the trace computation delivers a *sum* over the set of places $S \subset \Sigma_{\mathbb{K}}$.

Throughout this chapter, we only deal with the semi-local version of the trace formula in the Hilbert space context, as in [**71**]. In Chapter 4 we return to discuss, at length, the trace formula and the spectral realization, but in a different setting like that of [**74**], where we give a cohomological interpretation of the trace formula. In that setting we will have all the zeros appearing (not just those on the critical line), while the Riemann Hypothesis will become a question equivalent to a positivity statement that resembles more closely the positivity argument in Weil's proof for function fields.

The terms that appear in the *semi-local* trace formula are compared in §8 with Weil's distributional formulation of Riemann's explicit formula relating primes and zeros of the zeta function. We begin by recalling in §§8.1 and 8.2 the definition of L-functions with Grössencharakter for a global field $\mathbb{K}$ (the Riemann zeta function being a special case) and the distribution $\mathbf{D}_{\mathbb{K}}$ on $C_{\mathbb{K}}$ and the Weil principal values that enter in the explicit formula. In §8.3 we recall the properties of Fourier transform on $C_{\mathbb{K}}$. In §8.4 we show the explicit computation of the principal values comparing the Weil principal values with those considered earlier in the semi-local trace formula of §7.2. In particular, we discuss separately in §§8.4.1, 8.4.2, and 8.4.3 the case of a non-Archimedean place and the cases of a real and a complex Archimedean place. Finally, in §8.5, we give a reformulation of the Weil explicit formula in the form

$$\hat{h}(0) + \hat{h}(1) - \sum_{\chi \in \widehat{C_{\mathbb{K},1}}} \sum_{Z_{\tilde{\chi}}} \hat{h}(\tilde{\chi}, \rho) = \sum_v \int'_{\mathbb{K}_v^*} \frac{h(u^{-1})}{|1-u|} \, d^*u,$$

where the principal values match the corresponding terms that appear in the semi-local trace formula of §7.2.

We then discuss the spectral realization of [**71**] of the zeros of the Riemann zeta function and of L-functions with Grössencharakter. We consider the cokernel of the map $\mathfrak{E}$, extended to this adelic setting, as an isometry $\mathfrak{E} : L^2_\delta(X_{\mathbb{K}})_0 \to L^2_\delta(C_{\mathbb{K}})$ of the form

$$\mathfrak{E}(f)\,(g) = |g|^{1/2} \sum_{q \in \mathbb{K}^*} f(qg)$$

for $g \in C_{\mathbb{K}}$, where the domain $L^2_\delta(X_{\mathbb{K}})_0$ is the subspace defined by the conditions $f(0) = \hat{f}(0) = 0$ in the Hilbert space of square-integrable functions on $X_{\mathbb{Q}}$ with a weight δ controlling the decay condition. The cokernel $\mathcal{H}$ of $\mathfrak{E}$ carries an induced representation $\underline{\vartheta}_{\mathrm{m}}(g)$, $g \in C_{\mathbb{K}}$ and can be decomposed as a sum $\bigoplus_\chi \mathcal{H}_\chi$ along characters. The infinitesimal generator of the induced scaling action of $\mathbb{R}_+^*$ on $\mathcal{H}_\chi$ has as spectrum the set of zeros on the critical line of the corresponding L-function with Grössencharakter. The trace is then given by

$$\operatorname{Tr} \underline{\vartheta}_{\mathrm{m}}(h) = \sum_{\substack{L\left(\tilde{\chi}, \frac{1}{2}+\rho\right)=0 \\ \rho \in i\,\mathbb{R}/N^{\perp}}} \hat{h}(\tilde{\chi}, \rho),$$

for $\underline{\vartheta}_{\mathrm{m}}(h) = \int \underline{\vartheta}_{\mathrm{m}}(\gamma)\, h(\gamma)\, d^*\gamma$ with $h \in \mathcal{S}(C_{\mathbb{K}})$. The proof of [**71**] of the theorem on the spectral realization (Theorem 2.47) is given then in §9.3, after discussing in §9.1 a way to express L-functions as normalization factors of certain homogeneous distributions on $\mathbb{A}_{\mathbb{K}}$ and, in §9.2, the use of approximate units f_n in the Sobolev space $L^2_\delta(C_{\mathbb{K}})$.

In §10 we consider instead L-functions of varieties defined over number fields as in [**74**]. We concentrate on the Archimedean local factors of the completed L-function. These are suitable products of gamma functions with shifts and exponents that depend upon the Hodge numbers of the complex algebraic variety determined by the embedding $\mathbb{K} \hookrightarrow \mathbb{C}$ defining an Archimedean place. By analogy with the Riemann zeta function and L-functions with Grössencharakter discussed in the previous sections, we consider the counting function of zeros for L-functions of varieties. Conjecturally, the L-functions $L(H^m(X), s)$ satisfy a functional equation and have zeros located on the critical line $\Re(s) = (1+m)/2$. The average part of the counting function $N(E) = \langle N(E)\rangle + N_{\mathrm{osc}}(E)$ should then be expressed in terms of the Archimedean factors

$$\langle N(E)\rangle = \frac{1}{\pi} \sum_{v|\infty} \Im \log L_v \left(H^m(X), \frac{1+m}{2} + iE\right).$$

We show that this expression can be formulated in terms of a Lefschetz trace formula, much as in the case of the Riemann zeta function. After recalling in §10.1 the definition of the Archimedean local L-factors and the counting function in §10.2, we proceed by expressing their logarithmic derivatives in distributional form in terms of principal values in §10.3. We show in §§10.4 and 10.5 that this produces, respectively in the case of a complex or a real place, a trace formula for the logarithmic derivative of the corresponding local factor. In §§10.6 and 10.7 we formulate some more general questions such as the problem of a trace formula involving several places and the related spectral realization, and the problem of finding a global space (the analog in this higher-dimensional setting of the adèle class space) on which the geometric side of the expected trace formula lives. We conclude the chapter with §10.8, where we draw an analogy between the real mixed Hodge structures involved in the definition of the Archimedean local L-factors and their motivic Galois group and the category of flat equisingular connections used in Chapter 1 in the context of perturbative renormalization, and their differential Galois group $\mathbb{U}^*$.

2. Counting primes and the zeta function

We begin by recalling some very well known facts about the Riemann zeta function. We do not prove these classical statements here, but the interested reader can find plenty of material on this subject, for instance by looking at the very pleasant book

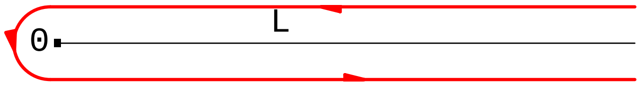


FIGURE 1. Contour L

by Edwards [**126**] in which Riemann's original paper, which we essentially follow below, is reproduced.

Riemann's seminal paper "On the number of primes less than a given magnitude" (in [**254**], cf. also [**126**]) established a remarkable relation between the distribution of prime numbers and the zeros of the zeta function given by the Euler product

$$\zeta(s) = \sum_{n \geq 1} n^{-s} = \prod_p (1 - p^{-s})^{-1}, \quad \Re(s) > 1, \tag{2.1}$$

where p ranges over the prime numbers.

The function defined by (2.1) has analytic continuation to the complex plane and is regular everywhere except for a simple pole at $s = 1$, where

$$\lim_{s \to 1} \zeta(s) - \frac{1}{s-1} = \gamma,$$

with γ the Euler constant. The analytic continuation is obtained in [**254**] using the Γ function

$$\Gamma(s)\, n^{-s} = \int_0^\infty e^{-nx} x^{s-1}\, dx,$$

which gives

$$\int_0^\infty \frac{x^{s-1}}{e^x - 1}\, dx = \Gamma(s) \sum_{n=1}^\infty n^{-s},$$

from which one obtains

$$\zeta(s) = \frac{\Gamma(1-s)}{2\pi i} \int_L \frac{(-x)^s}{e^x - 1} \frac{dx}{x},$$

where

$$(-x)^s = e^{s \log(-x)}, \quad \Im \log(z) \in (-\pi, \pi), \quad \forall z \notin -\mathbb{R}_+$$

and the integration is along the oriented path L from $+\infty$ to $i\epsilon$ in $\mathbb{R} + i\epsilon$, then counterclockwise around the origin in $\mathbb{C}$ and then back from $-i\epsilon$ to $+\infty$ in $\mathbb{R} - i\epsilon$.

Using Cauchy's residue formula Riemann proves the functional equation

$$\zeta(1-s) = 2(2\pi)^{-s} \cos(s\pi/2) \Gamma(s) \zeta(s), \tag{2.2}$$

which he writes in more symmetric form as the symmetry

$$\zeta^*(s) = \zeta^*(1-s)\,, \tag{2.3}$$

of the "complete" zeta function

$$\zeta^*(s) = \Gamma(s/2) \pi^{-s/2} \zeta(s)\,. \tag{2.4}$$

After writing ζ^* as the Mellin transform of Jacobi's theta function, $\vartheta(x) = \sum_1^\infty e^{-n^2\pi x}$,

$$\zeta^*(s) = \int_0^\infty \vartheta(x)\, x^{\frac{s}{2}-1}\, dx\,, \tag{2.5}$$

which gives another proof of the functional equation (2.3), Riemann then introduces the ξ function, which he defines by

$$\xi(t) = \frac{s(s-1)}{2}\Gamma(s/2)\pi^{-s/2}\zeta(s)\,, \quad s = \frac{1}{2} + it \tag{2.6}$$

The function $\xi(t)$ is entire, even and real-valued for real t,

$$\xi\,(t) \in \mathbb{R}, \quad \forall t \in \mathbb{R}.$$

It was stated in Riemann's paper (and rigorously proved by Hadamard) that the ξ function can also be written in the form

$$\xi(t) = \xi(0) \prod_\alpha \left(1 - \frac{t}{\alpha}\right), \tag{2.7}$$

with the product ranging over the zeros $\{\alpha \in \mathbb{C} : \xi(\alpha) = 0\}$. The Riemann Hypothesis, still unproved at the time of this writing, states that these zeros are all real.

The zeta function $\zeta(s)$ has zeros at all the negative even integers $s = -2, -4, \ldots, -2n, \ldots$. These are usually referred to as the *trivial zeros*. The hard problem is the problem of locating the remaining *nontrivial zeros*. From (2.6) they are of the form

$$\rho = \frac{1}{2} + i\,\alpha\,, \quad \xi(\alpha) = 0$$

and the Riemann Hypothesis states that they all fall on the line $\Re(z) = 1/2$.

The explicit relation between the zeros of zeta and prime numbers is expressed through the *explicit formulas*. In its original form due to Riemann the explicit formula expresses the *prime counting function*

$$\pi(x) = \#\{p \text{ prime } : p \le x\}. \tag{2.8}$$

in terms of the integral logarithm

$$\mathrm{Li}(x) = \int_0^x \frac{du}{\log(u)} \sim \sum_k (k-1)! \frac{x}{\log(x)^k}. \tag{2.9}$$

(which Gauss had recognized numerically as giving the leading behaviour of $\pi(x)$) and the non-trivial zeros ρ of zeta in the form

$$\pi'(x) = \mathrm{Li}(x) - \sum_\rho \mathrm{Li}(x^\rho) + \int_x^\infty \frac{dt}{t(t^2-1)\log t} + \log \xi(0), \tag{2.10}$$

where one lets

$$\pi'(x) = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) + \cdots + \frac{1}{n}\pi(x^{1/n}) + \cdots \tag{2.11}$$

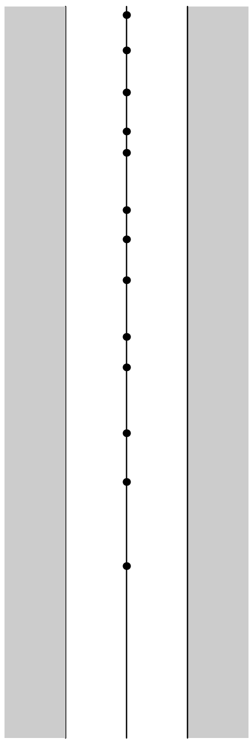


FIGURE 2. Zeros of Zeta

The *prime number theorem* asserts that the counting function has the asymptotic behavior

$$\pi(x) \sim \frac{x}{\log(x)} \tag{2.12}$$

The Riemann Hypothesis mentioned above is equivalent to the statement that the remainder in

$$\pi(x) = \mathrm{Li}(x) + R(x) \tag{2.13}$$

has size of the order of

$$|R(x)| = O(\sqrt{x}\log(x)). \tag{2.14}$$

Note that the main virtue of (2.10) is not to give an explicit formula for the prime counting function. Such an explicit formula exists [**201**] and is given (for $n > 5$) by

$$\pi(n) = 2 + \sum_{k=5}^{n} \frac{e^{2\pi i \Gamma(k)/k} - 1}{e^{-2\pi i/k} - 1}. \tag{2.15}$$

It is clear though, that because of the presence of the large $\Gamma(k) = (k-1)!$ in the exponential $e^{2\pi i \Gamma(k)/k}$, the formula (2.15) is of little use in extracting an asymptotic estimate like the desired (2.14).

There are many variants of the explicit formula (2.10). For instance let $\Lambda(n)$ be the von Mangoldt function

$$\Lambda(n) = \begin{cases} \log p & n = p^k\,,\ p \text{ prime}\,,\ k \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

and $\psi(x)$ be the Chebyshev–von Mangoldt function,

$$\psi(x) = \sum_{n \le x} \Lambda(n)\,, \tag{2.17}$$

i.e. the step function that starts at zero and jumps by $\log p$ at each prime power p^k. The value of $\psi(x)$ at a point of discontinuity is *defined* as the average

$$\psi(x) = \frac{\psi(x+0) + \psi(x-0)}{2}\,. \tag{2.18}$$

Then one instance of the explicit formula is ([**239**] Theorem 3.8)

$$\psi(x) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \log(2\pi) - \frac{1}{2}\log(1 - x^{-2})\,, \quad \forall x > 1\,. \tag{2.19}$$

Here the sum is over all nontrivial zeros of $\zeta(s)$ and is understood as the limit

$$\lim_{E \to \infty} \sum_{|\Im(\rho)| < E} \frac{x^\rho}{\rho}.$$

All of these formulas are special cases of Weil's explicit formula which is best stated in the framework of test functions on the idèle class group. We shall recall this formula below in §8 and one of the main results of this chapter, besides providing a spectral interpretation of the zeros of the zeta function, is to give a geometric set-up in which Weil's explicit formula becomes a trace formula.

In his paper Riemann gives an asymptotic estimate for the counting of the nontrivial zeros of zeta in a certain region, namely the function

$$N(E) = \#\{\rho \,|\, \zeta(\rho) = 0, \ \text{ and } \ 0 < \Im(\rho) \le E\}. \tag{2.20}$$

As a preliminary result to the study of the function $\pi(x)$, Riemann states that this counting function has the asymptotic behavior

$$N(E) \sim \frac{E}{2\pi} \log \frac{E}{2\pi} - \frac{E}{2\pi}\,. \tag{2.21}$$

This was rigorously proved by von Mangoldt, and can be seen by considering the integral

$$2\,(N(E) - 1) = \frac{1}{2\pi i} \int_{\partial R} \frac{d\zeta^*(s)}{\zeta^*(s)} \tag{2.22}$$

counterclockwise along the boundary of the rectangle $R = \{s : -\epsilon \le \Re(s) \le 1 + \epsilon, -E \le \Im(s) \le E\}$ (cf. Figure 3). Using the expression (2.4), one obtains (cf. e.g. [**126**], §6.6)

$$N(E) = \frac{\theta(E)}{\pi} + 1 + \frac{1}{\pi} \Im \int_C \frac{\zeta'(s)}{\zeta(s)} ds, \tag{2.23}$$

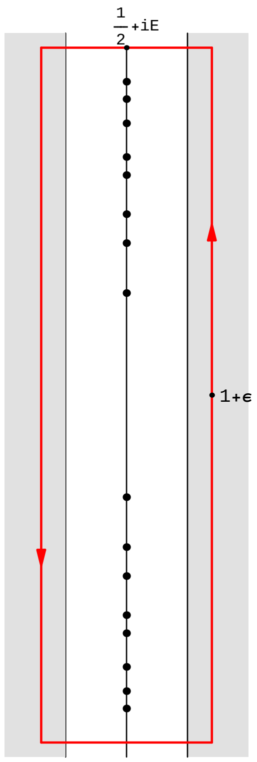


FIGURE 3. The rectangle R and the path C

where C is the part of ∂R from $1+\epsilon$ to $1/2+iE$ (cf. Figure 3) and $\theta(E)$ is the Riemann–Siegel angular function, given by

(2.24) $$\theta(E) = -\frac{E}{2}\log\pi + \Im\log\Gamma\left(\frac{1}{4}+i\frac{E}{2}\right),$$

with $\log\Gamma(s)$, for $\Re(s) > 0$ the branch of the log which is real for s real. This represents the opposite of the phase of zeta on the critical line inasmuch as the function $Z(t)$,

$$\zeta\left(\frac{1}{2}+it\right) = e^{-i\theta(t)}Z(t),$$

is real-valued for $t \in \mathbb{R}$. For this function $\theta(t)$ one has (cf. e.g. [**126**] §6.5)

$$\theta(t) = \frac{t}{2}\log\frac{t}{2\pi} - \frac{t}{2} - \frac{\pi}{8} + \frac{1}{48t} + \frac{7}{5760t^3} + \cdots$$

from which the estimate for $N(E)$ follows, provided one controls the size of the error term

(2.25) $$N_{\mathrm{osc}}(E) = \frac{1}{\pi}\Im\int_C \frac{\zeta'(s)}{\zeta(s)}ds = \frac{1}{\pi}\Im\log\zeta\left(\frac{1}{2}+iE\right),$$

where we assume that E is not the imaginary part of a zero, while the branch of the logarithm is uniquely defined by (2.25) and is 0 at $iE+\infty$. The main contribution to the error term $N_{\mathrm{osc}}(E)$ comes from the integral in the portion of the contour C

from $1+iE$ to $\frac{1}{2}+iE$ and one can show (cf. [**239**] Theorem 4.7) that

$$N_{\mathrm{osc}}(E) = O(\log E)\,.$$

3. Classical and quantum mechanics of zeta

The above result of Riemann on the estimate for $N(E)$ as $E \to \infty$, which we discussed in §2 (cf. [**254**], [**126**]) can be formulated as the statement that the function $N(E)$ of (2.20) has a decomposition of the form

$$N(E) = \langle N(E)\rangle + N_{\mathrm{osc}}(E)\,, \tag{2.26}$$

where the oscillatory part $N_{\mathrm{osc}}(E)$ is of the order of $O(\log E)$ and is given by the expression

$$N_{\mathrm{osc}}(E) = \frac{1}{\pi}\Im \log\zeta\left(\frac{1}{2}+iE\right)\,, \tag{2.27}$$

while the average part

$$\langle N(E)\rangle = 1 + \frac{\theta(E)}{\pi}\,, \tag{2.28}$$

has an asymptotic expansion of the form

$$\langle N(E)\rangle = \frac{E}{2\pi}\left(\log\frac{E}{2\pi} - 1\right) + \frac{7}{8} + o(1)\,. \tag{2.29}$$

One has the following equivalent expressions for $\langle N(E)\rangle$.

LEMMA 2.1. *The function $\langle N(E)\rangle$ satisfies*

$$\langle N(E)\rangle = 1 - \frac{E}{2\pi}\log\pi + \frac{1}{\pi}\Im\log\Gamma\left(\frac{1}{4}+i\frac{E}{2}\right)\,, \tag{2.30}$$

and

$$\langle N(E)\rangle = 1 + \frac{1}{2\pi}\int_0^E\left(-\log\pi + \Re\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}+i\frac{s}{2}\right)ds\right)\,. \tag{2.31}$$

PROOF. The first equality follows from (2.28) and (2.24). The second follows from

$$d\Im\log\Gamma\left(\frac{1}{4}+i\frac{s}{2}\right) = \frac{1}{2i}d\left(\log\Gamma\left(\frac{1}{4}+i\frac{s}{2}\right) - \log\Gamma\left(\frac{1}{4}-i\frac{s}{2}\right)\right)$$
$$= \frac{1}{2}\Re\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}+i\frac{s}{2}\right)ds$$

while $\Im\log\Gamma\left(\frac{1}{4}+i\frac{s}{2}\right) = 0$ for $s=0$ by the choice of the branch of the log. ∎

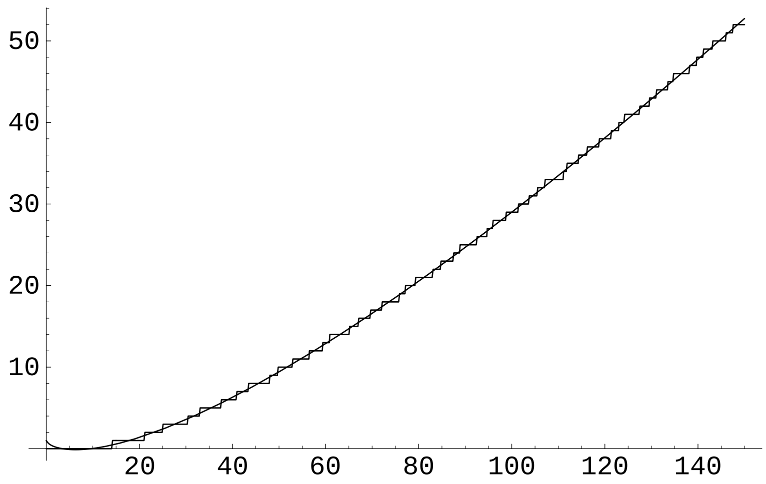


FIGURE 4. The number of zeros $N(E)$ and its approximation by $\langle N(E)\rangle$.

3.1. Spectral lines and the Riemann flow.

It was observed by Berry in [**18**] that there is an interesting similarity between the expression in (2.26) and the semi-classical formula for the number of eigenvalues

$$N(E) = \#\text{ Eigenvalues of } H \text{ in } [0, E] \tag{2.32}$$

of a Hamiltonian operator H obtained by "quantizing" a classical dynamical system given by a symplectic phase space X and a Hamiltonian function h. In this setting, the average part $\langle N(E)\rangle$ is computed as a volume in phase space, namely

$$\langle N(E)\rangle := \text{Symplectic Volume } \{x \in X \mid h(x) \leq E\}, \tag{2.33}$$

while the oscillatory part is given by

$$N_{\rm osc}(E) = N(E) - \langle N(E)\rangle. \tag{2.34}$$

In the Hamiltonian case $N_{\rm osc}(E)$ is the same as in random matrix theory, with statistics dictated by the symmetries of the system.
This analogy led to the idea (cf. [**18**]) of the existence of a certain (hypothetical) "Riemann flow". This would be a Hamiltonian system (X, h) which would give the formula (2.29) from (2.33). It should also match the heuristic formula, obtained by naively applying the function $\frac{1}{\pi}\Im \log$ to the Euler product formula using (2.27),

$$N_{\rm osc}(E) \sim \frac{-1}{\pi} \sum_p \sum_{m=1}^{\infty} \frac{1}{m} \frac{1}{p^{m/2}} \sin\left(m\, E \,\log\, p\right) \tag{2.35}$$

with the semi-classical formula for the oscillatory part $N_{\rm osc}(E)$ in the Hamiltonian case. The latter is of the form

$$N_{\rm osc}(E) \sim \frac{1}{\pi} \sum_{\gamma_p} \sum_{m=1}^{\infty} \frac{1}{m} \frac{1}{2\sinh\left(\frac{m\lambda_p}{2}\right)} \sin(m\, E\, T_\gamma^\#), \tag{2.36}$$

where the γ_p are the primitive periodic orbits of the system, $T_\gamma^\#$ is the period of the primitive orbit γ_p, the label m corresponds to the number of traversals of this orbit, and the $\pm\lambda_p$ are the corresponding instability exponents.

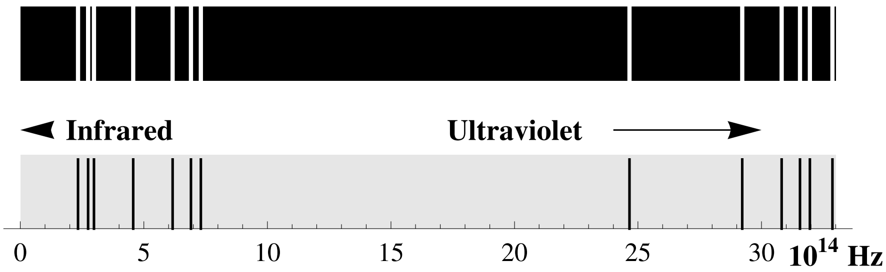


FIGURE 5. Emission and absorption frequency spectra of Hydrogen.

One can obtain some interesting information on the properties of such a conjectural Riemann flow by comparing the formulas (2.35) and (2.36). However, while there seems to be a clear meaning for the various terms in (2.35) in this analogy, there is a serious problem that stems from the overall minus sign in front of (2.35), which cannot be accounted for in (2.36) even if one tries to fiddle with Maslov indices.

A solution to this sign problem was given by Connes in [**71**]. A physical interpretation is the following. In quantum physics the observed spectra are of two different kinds (cf. Figure 5). Emission spectra appear as bright lines on a dark background and are generated by emitted photons of the corresponding frequency. Absorption spectra appear as dark lines on a clear background, and look like the "negatives" of emission spectra. There is a fundamental difference between the two kinds of spectra due to the presence of the continuous spectrum of white light which is the background of all absorption spectra. The dark lines correspond to lacunae in the continuous spectrum due to the absorption of those photons that have the right frequency to induce atomic transitions during their trajectory from the source of white light to the photographic plate. While the excited atoms re-emit photons of the same frequency, the directions of these re-emitted photons are random and this lack of coherence is responsible for the dark "absorption" lines in the spectrum. These dark lines would not be seen if they were sharp, but they have a natural width due to several effects, the simplest being the lifetime of the excited state and the presence of natural infrared and ultraviolet cutoffs. Mathematically speaking, as we discuss in the following, talking about an absorption spectrum is akin to considering a cokernel and we will see that the presence of cutoffs also plays an important role in this formulation.

3.2. Symplectic volume and the scaling Hamiltonian.

We begin by computing the symplectic volume (2.33) for the classical mechanical system with phase space $T^*\mathbb{R}$, the cotangent space of the real line, and classical Hamiltonian $h(q,p)$ that generates the group of scaling transformations. More precisely, we consider the product $\mathbb{R} \times \hat{\mathbb{R}}$ of the additive group $\mathbb{R}$ by its dual group $\hat{\mathbb{R}}$ and we use the bicharacter

$$e_{\mathbb{R}}(p,q) = \exp(-2\pi i pq) \tag{2.37}$$

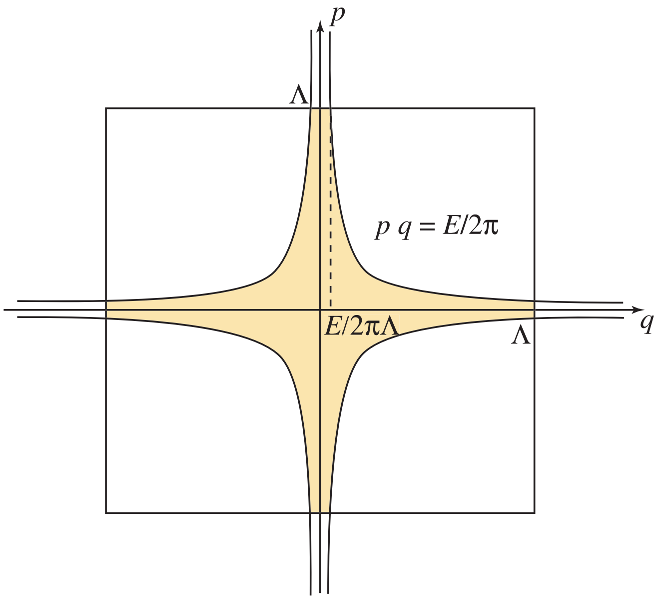


FIGURE 6. The subset B of the (q, p) plane

to identify $\hat{\mathbb{R}}$ with $\mathbb{R}$. In this notation $\frac{\partial}{\partial q}$ corresponds to $-2\pi ip$ and the generator $iq\frac{\partial}{\partial q}$ of the scaling transformations corresponds to $2\pi qp$. Thus, the Hamiltonian $h(q, p)$ is of the form

$$h(q, p) = 2\pi qp\,. \tag{2.38}$$

The canonical symplectic form

$$\omega = dp \wedge dq \tag{2.39}$$

is the product of the Haar measure on $\mathbb{R}$ by the dual Haar measure on $\hat{\mathbb{R}}$.
We let X denote the quotient of $\mathbb{R} \times \hat{\mathbb{R}}$ by the discrete symmetry $(p, q) \mapsto (-p, -q)$. The Hamiltonian $h(q, p)$ is not positive, but this is in line with the symmetry $\frac{1}{2} + iE \mapsto \frac{1}{2} - iE$ on zeros of ζ. This symmetry shows that we should compare

$$2N(E) = \#\{\rho \mid \zeta^*(\rho) = 0\,,\ |\Im\rho| \leq E\} \tag{2.40}$$

with the symplectic volume of $|h| \leq E$ in (2.33), i.e. with the symplectic volume of the region satisfying

$$|qp| \leq \frac{E}{2\pi}\,.$$

We introduce a cutoff Λ both in the infrared and in the ultraviolet by imposing the conditions

$$|q| \leq \Lambda\,, \quad |p| \leq \Lambda. \tag{2.41}$$

We then compute the symplectic volume $v(E)$ of the subset B of X:

$$B = \{(q, p)\ :\ |h| \leq E\,, \quad |q| \leq \Lambda\,, \quad |p| \leq \Lambda\}. \tag{2.42}$$

By construction B is the image in X of the union of four regions congruent to its intersection B_+ with the first quadrant, as shown in Figure 6.

Thus, the symplectic volume $v(E)$ of the region B in X is given by twice the symplectic volume of B_+,

$$v(E) = 2\,\mathrm{Vol}(B_+).$$

Notice that B_+ is the union of the rectangle with sides $\frac{E}{2\pi\Lambda}$ and Λ and of the region below the graph, in the range from $q = \frac{E}{2\pi\Lambda}$ to $q = \Lambda$, of the hyperbola $2\pi pq = E$. We assume $\Lambda > (\frac{E}{2\pi})^{1/2}$. Thus, we have

$$\int_{B_+} \omega = \frac{E}{2\pi} + \int_{E/2\pi\Lambda}^{\Lambda} \frac{E}{2\pi q} dq\,. \tag{2.43}$$

This gives

$$\mathrm{Vol}(B_+) = \frac{E}{2\pi} \times 2\log\Lambda - \frac{E}{2\pi}\left(\log\frac{E}{2\pi} - 1\right). \tag{2.44}$$

In this formula we see the overall term $\langle N(E)\rangle$ of (2.28) appearing with a *minus* sign, as

$$-\frac{E}{2\pi}\left(\log\frac{E}{2\pi} - 1\right). \tag{2.45}$$

In §5 below we refine this semi-classical approximation and obtain in Proposition 2.17 exactly $-\langle N(E)\rangle$ of (2.28) (and in particular the additional term $\frac{7}{8}$) as a quantum correction.

The terms in $\log\Lambda$ are of the form $\frac{E}{2\pi} \times 2\log\Lambda$. This contributes an overall factor of 2 to $v(E)$. Now $2 \times \frac{E}{2\pi} \times 2\log\Lambda$ is the symplectic volume of the cutoff region

$$W(E,\Lambda) = \{(\lambda,\theta) \in \mathbb{R}_+^* \times \mathbb{R} \mid \lambda \in [\Lambda^{-1},\Lambda]\,, \quad |\theta| \le E\} \tag{2.46}$$

in the product of the multiplicative group $\mathbb{R}_+^*$ by its dual group $\hat{\mathbb{R}}_+^* \simeq \mathbb{R}$. Here we identify $\hat{\mathbb{R}}_+^*$ with $\mathbb{R}$ using the bicharacter

$$\langle \lambda, \theta\rangle = \lambda^{i\theta}. \tag{2.47}$$

In fact, the product of the Haar measure on $\mathbb{R}_+^*$ by the dual Haar measure on $\mathbb{R}$ is given by

$$\frac{1}{2\pi}\frac{d\lambda}{\lambda} \times d\theta. \tag{2.48}$$

Thus, we can see clearly, already at this rough semi-classical level, that the symplectic volume $v(E)$ for the scaling flow on X is given by the difference

$$v(E) = \mathrm{Vol}(W(E,\Lambda)) - 2\langle N(E)\rangle, \tag{2.49}$$

where $2\langle N(E)\rangle$ is the average part of the counting function $2N(E)$ of (2.40).

This suggests that the scaling flow on X can be regarded as a first approximation to a classical system whose quantization yields a spectral interpretation of the zeros of zeta, not as an emission spectrum but as an absorption spectrum, with the zeros labeling the missing spectral lines, while $\mathrm{Vol}(W(E,\Lambda))$ accounts for the white light background.

Besides the first difficulty, which is to quantize the above system and replace the simple area computation by the counting of the modes of the quantum Hamiltonian, the main difficulty comes from accounting for all the additional terms which enter in the oscillatory part of $N(E)$ as given in (2.27). We shall start with the first question, i.e. the counting of modes of the Hamiltonian of the quantum system.

3.3. Quantum system and prolate functions.

Passing from the classical to the quantum mechanical system replaces integration of the symplectic form over a region of the phase space by a counting of quantum states within a given energy bound.
The classical system (X, h) described in §3.2 above is easy to quantize. The quantum mechanical Hilbert space is simply given by

$$\mathcal{H} = L^2(\mathbb{R})^{\mathrm{even}}, \tag{2.50}$$

that is, the space of square-integrable *even* functions $f(q)$. Requiring the functions to be even simply reflects the discrete symmetry $(p, q) \mapsto (-p, -q)$ defining X in the classical setting.

The Hamiltonian H generates the group of scaling transformations, which is given by the natural representation ϑ_{a} of $\mathbb{R}_+^*$ on $\mathcal{H}$, of the form

$$(\vartheta_{\mathrm{a}}(\lambda)\, f)(q) = f(\lambda^{-1} q). \tag{2.51}$$

This is unitary after multiplication by $|\lambda|^{-1/2}$ so that the representation

$$\lambda \mapsto |\lambda|^{-1/2} \vartheta_{\mathrm{a}}(\lambda) \tag{2.52}$$

is unitary. For $h \in \mathcal{C}_c^\infty(\mathbb{R}_+^*)$, we define corresponding operators

$$\vartheta_{\mathrm{a}}(h) = \int_{\mathbb{R}_+^*} h(\lambda) \vartheta_{\mathrm{a}}(\lambda)\, d^*\lambda \tag{2.53}$$

with $d^*\lambda = \frac{d\lambda}{\lambda}$ the multiplicative Haar measure on $\mathbb{R}_+^*$.
To the infrared cutoff $|q| \leq \Lambda$ there corresponds the orthogonal projection P_Λ onto the subspace

$$P_\Lambda = \{ f \in L^2(\mathbb{R})^{\mathrm{even}} \,|\, f(q) = 0,\ \forall q \ \text{with}\ |q| > \Lambda \}. \tag{2.54}$$

To define the ultraviolet cutoff $|p| \leq \Lambda$ we use the Fourier transform, which is defined as follows for locally compact abelian groups:

DEFINITION 2.2. *Let A, B be a pair of locally compact abelian groups with Haar measures da, db. Let $\alpha(a, b)$ be a bicharacter which gives an isomorphism $B \sim \hat{A}$ of B with the Pontrjagin dual of A. Then the Fourier transform $\mathbf{F}_\alpha$ is defined by*

$$\mathbf{F}_\alpha(f)(b) := \int \alpha(a, b)\, f(a)\, da\,. \tag{2.55}$$

The ultraviolet cutoff is given by the orthogonal projection $\widehat{P}_\Lambda$ given by the conjugate

$$\widehat{P}_\Lambda = \mathbf{F}_{e_{\mathbb{R}}} P_\Lambda \mathbf{F}_{e_{\mathbb{R}}}^{-1}. \tag{2.56}$$

Here $\mathbf{F}_{e_{\mathbb{R}}}$ is the Fourier transform associated to the bicharacter $e_{\mathbb{R}}$ introduced above in (2.37).
The first difficulty one encounters, which is inherent in the quantum system, is the fact that the two cutoffs P_Λ and $\widehat{P}_\Lambda$ do not commute, hence one cannot just intersect their ranges in order to perform both cutoffs at once. Indeed, no nonzero function on $\mathbb{R}$ can have the property that both the function and its Fourier transform have compact support.

The relative position of the two projections (2.54) and (2.56) was first analyzed in the work of Landau, Pollak and Slepian (cf. [**245**], [**196**], [**197**]) in the early sixties, motivated by problems of electrical engineering.
The result has since then been considerably refined. To explain the main point, let us first remind the reader of well-known facts on the geometry of pairs of projections.

LEMMA 2.3. *Giving a pair of orthogonal projections P_i, $i = 1, 2$ on a Hilbert space $\mathcal{H}$ is the same as giving a unitary representation of the dihedral group $\Gamma = \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$, the free product of two copies of the group $\mathbb{Z}/2\mathbb{Z}$. These irreducible unitary representations are parameterized by an angle $\alpha \in [0, \pi/2]$.*

PROOF. To the generators U_i of Γ there correspond the operators $2P_i - 1$. The group Γ is the semidirect product of the subgroup generated by $U = U_1 U_2$ by the group $\mathbb{Z}/2\mathbb{Z}$, acting by $U \mapsto U^{-1}$. Irreducible unitary representations of Γ are parameterized by an angle $\alpha \in [0, \pi/2]$, where the corresponding orthogonal projections P_i are associated to the one-dimensional subspaces $y = 0$ and $y = x \tan(\alpha)$ in the Euclidean (x, y)-plane. In particular these representations are at most two-dimensional. The projections P_i are given by

$$P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; P_2 = \begin{pmatrix} \cos^2\alpha & \cos\alpha\,\sin\alpha \\ \cos\alpha\,\sin\alpha & \sin^2\alpha \end{pmatrix}. \tag{2.57}$$

A general unitary representation is characterized by an operator α whose value is the angle α in the irreducible case. It is uniquely defined by the condition $0 \leq \alpha \leq \pi/2$ and the equality

$$\sin(\alpha) = |P_1 - P_2| \tag{2.58}$$

and commutes with P_i. Equivalently one has

$$P_1\, P_2\, P_1 = \cos^2(\alpha) P_1\,. \tag{2.59}$$

$\square$

What the authors of [**245**], [**196**], [**197**] did is to analyze the relative position of the projections P_Λ and $\widehat{P}_\Lambda$ for $\Lambda \to \infty$. In concrete problems of electrical engineering, this makes it possible to account for signals which, for all practical purposes, have finite support both in the time variable and the dual frequency variable. This had an important role in the early developments of laser technology.
The basic tool employed in [**245**], [**196**], and [**197**] is the existence of a second-order differential operator $\mathbf{W}_\Lambda$ on $\mathbb{R}$, which commutes with both P_Λ and $\widehat{P}_\Lambda$. This operator is of the form

$$(\mathbf{W}_\Lambda \psi)(q) = -\partial((\Lambda^2 - q^2)\partial)\,\psi(q) + (2\pi\Lambda q)^2\,\psi(q). \tag{2.60}$$

Here ∂ denotes the ordinary differentiation in one variable. To be more precise, (2.60) defines a symmetric operator with natural domain the Schwartz space. Recall (cf. [**183**]) that by definition the deficiency indices $n_\pm(T)$ of a symmetric operator T are the dimensions of $\ker(T^* \pm i)$. One can show that the operator $\mathbf{W}_\Lambda$ has both deficiency indices equal to 4 and admits a unique selfadjoint extension which commutes with the dihedral group Γ associated to the projections P_Λ and $\widehat{P}_\Lambda$. We use the same notation $\mathbf{W}_\Lambda$ to denote this selfadjoint operator. It commutes with the Fourier transform $\mathbf{F}_{e_\mathbb{R}}$.

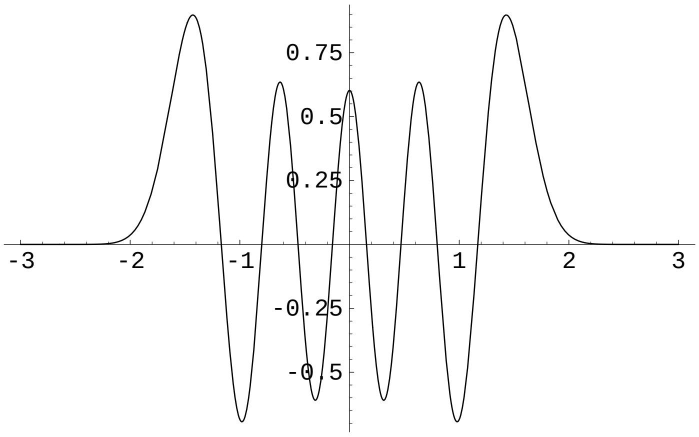


FIGURE 7. The function ψ_8 for $\Lambda = 3$.

If one restricts the variable to the range $[-\Lambda, \Lambda]$, the operator $\mathbf{W}_\Lambda$ has discrete simple spectrum. Its properties were studied long ago. It appears in the factorization of the Helmholtz equation

$$\Delta\,\psi + k^2\,\psi = 0 \tag{2.61}$$

in one of the few separable coordinate systems in Euclidean 3-space, called *prolate spheroidal coordinates*. The eigenvalues $\chi_n(\Lambda), n \geq 0$ of $\mathbf{W}_\Lambda$, listed in increasing order $\chi_{n+1} > \chi_n$, are simple and positive. The corresponding eigenfunctions ψ_n are called the *prolate spheroidal wave functions*. Since the product $P_\Lambda \widehat{P}_\Lambda P_\Lambda$ commutes with $\mathbf{W}_\Lambda$, they are eigenfunctions of $P_\Lambda \widehat{P}_\Lambda P_\Lambda$. By the results of [**245**], [**196**] and [**197**], the eigenvalues λ_n of the operator $P_\Lambda \widehat{P}_\Lambda P_\Lambda$ are simple. Moreover, if we let $\lambda_0 > \lambda_1 > \cdots > \lambda_n > \cdots$ be their list in decreasing order, one has

$$P_\Lambda \widehat{P}_\Lambda P_\Lambda\,\psi_n = \lambda_n \psi_n. \tag{2.62}$$

Thus, the prolate spheroidal wave functions ψ_n are the eigenvectors of $P_\Lambda \widehat{P}_\Lambda P_\Lambda$ with non-zero eigenvalues.

A lot is known about the prolate spheroidal wave functions ψ_n. In particular, one can take them to be real-valued. They are even for n even and odd for n odd, and they have exactly n zeros in the interval $[-\Lambda, \Lambda]$. We normalize them to have L^2 norm equal to 1. We are only interested in the ψ_n for even values of n, since in our quantum mechanical Hilbert space we restrict to even functions. For $\Lambda \to \infty$ the function ψ_n converges to the Hermite-Weber function Ω_n of order n (cf. [**271**]).
The eigenvalues λ_n behave qualitatively in the following manner, as a function of the cutoff. They stay very close to the value one, $\lambda_n \sim 1$, until n falls in an interval I of size $\sim \log \Lambda$ around the value $4\Lambda^2$. Their behavior in this interval is governed by the relation (cf. [**272**], [**198**])

$$\lambda_n = (1 + e^{\pi\delta})^{-1}, \tag{2.63}$$

where δ is the solution of smallest absolute value of the equation

$$(n + 1/2)\pi = 4\pi\Lambda^2 + \delta \log(8\pi\Lambda^2) - \delta(\log(|\delta/2|) - 1). \tag{2.64}$$

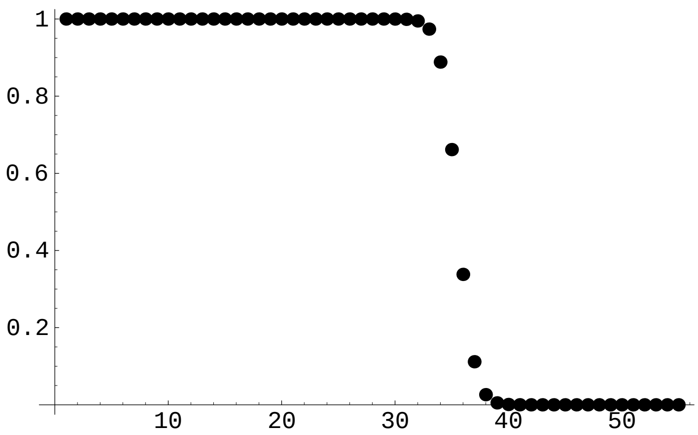


FIGURE 8. The eigenvalues λ_n for $\Lambda = 3$

Beyond the interval I the eigenvalues λ_n tend very rapidly to zero. A plot is shown in Figure 8.

For the pair of projections P_Λ and $\widehat{P}_\Lambda$, one obtains in this way the eigenvalues α_n of the angle operator α of Lemma 2.3 by (2.59), i.e.

$$\cos^2(\alpha_n) = \lambda_n. \tag{2.65}$$

This shows that the angle α_n is essentially 0 up to an interval I of size $\sim \log \Lambda$ around the value $n \sim 4\Lambda^2$. It then shifts from 0 to $\frac{\pi}{2}$ and stays essentially equal to $\frac{\pi}{2}$ for all larger values of n. This clearly shows how to impose both ultraviolet and infrared cutoffs by restricting to the subspace B_Λ of $\mathcal{H}$ spanned by the ψ_{2n}, for $2n \leq 4\Lambda^2$. No special care is needed in order to define precisely the upper bound, since one can show that, for all n in the interval I considered above and for all $h \in \mathcal{C}_c^\infty(\mathbb{R}_+^*)$, one has

$$\|\vartheta_{\mathrm{a}}(h)\,\psi_n\| = O(\Lambda^{-\rho}) \tag{2.66}$$

for some $\rho > 0$, and with $\vartheta_{\mathrm{a}}(h)$ as in (2.53). We let Q_Λ denote the orthogonal projection onto $B_\Lambda \subset \mathcal{H}$.

We now come to the problem of counting the number of states of the quantum mechanical system, with an energy bound $|H| \leq E$ on the Hamiltonian H (i.e. the generator of scaling transformations). The Hamiltonian H is the generator of the scaling action of $\mathbb{R}_+^*$ on $\mathcal{H}$ and we identify, as before, the dual group of $\mathbb{R}_+^*$ with $\mathbb{R}$ using the bicharacter (2.47).

LEMMA 2.4. *Let N_E denote the spectral projection, of the scaling action of $\mathbb{R}_+^*$ on $\mathcal{H}$, associated to the interval $[-E, E]$ in the dual group $\mathbb{R}$ of $\mathbb{R}_+^*$. It is given by*

$$N_E = \vartheta_{\mathrm{a}}(h_E), \quad \textit{with} \;\; h_E(u) = |u|^{-1/2}\frac{1}{2\pi}\int_{-E}^{E} |u|^{is} ds\,, \tag{2.67}$$

with $\vartheta_{\mathrm{a}}(h_E)$ defined in (2.53).

PROOF. As we saw above, (2.52) gives a unitary action of $\mathbb{R}_+^*$ on $\mathcal{H}$, so that (2.67) follows from the identity

$$\int \lambda^{it}\, k(\lambda)\, d^*\lambda = \mathbf{1}_{[-E,E]}(t)$$

for the function $k(\lambda) = \frac{1}{2\pi}\int_{-E}^{E} \lambda^{is} ds$. ☐

We can now formulate the problem of counting quantum states of the scaling flow in the following way.

REMARK 2.5. Counting the number of quantum states of the Hamiltonian H subject to the constraint $|H| \leq E$ amounts to computing the dimension of the near intersection of the projections Q_Λ and N_E. This is given by $\mathrm{Tr}(Q_\Lambda N_E)$.

In order to compute $\mathrm{Tr}(Q_\Lambda N_E)$ for large Λ, we can use the analysis of the angle operator between P_Λ and $\widehat{P}_\Lambda$ described above together with (2.66) and replace Q_Λ by

$$R_\Lambda = \widehat{P}_\Lambda\, P_\Lambda\,, \tag{2.68}$$

which introduces an error term of the order of $O(\Lambda^{-\rho}\log\Lambda)$. Thus, in order to count the number of quantum states in the near intersection of the two projections Q_Λ and N_E, we just need to compute

$$\mathrm{Tr}(Q_\Lambda N_E) \sim \mathrm{Tr}(R_\Lambda\, \vartheta_{\mathrm{a}}(h_E))\,. \tag{2.69}$$

It thus remains to compute $\mathrm{Tr}\,(R_\Lambda\, \vartheta_{\mathrm{a}}(h))$. We then apply the following result ([**71**], Theorem 3), which is in fact valid for any local field and is a special case of Theorem 2.36 which will be proved below, in §7.2.

THEOREM 2.6. *Let $h \in \mathcal{S}(\mathbb{R}^*)$ be a function with compact support, and $\vartheta_{\mathrm{a}}(h)$ be given by* (2.53)*. Then $R_\Lambda\, \vartheta_{\mathrm{a}}(h)$ is a trace-class operator. Moreover, for $\Lambda \to \infty$ one has*

$$\mathrm{Tr}\,(R_\Lambda\, \vartheta_{\mathrm{a}}(h)) = 2h(1)\log\Lambda + \int' \frac{h(u^{-1})}{|1-u|}\, d^*u + o(1). \tag{2.70}$$

We want to apply this formula to the function $h = h_E$ of (2.67). Notice that this function does not have compact support, so in principle the result does not apply directly. We shall take care of this technical point in §5.1 below. In (2.70) the integral is singular at $u = 1$ and Theorem 2.6 selects a specific principal value, denoted as $\int'$, which we now discuss in detail in §4 below.

4. Principal values from the local trace formula

The counting of the quantum states of the scaling flow makes sense for any local field K, i.e. any locally compact non-discrete field K. We shall prove below that its computation singles out a preferred extension of the distribution d^*u at $u = 0$. Here we shall first give a heuristic explanation for the local trace formula (Theorem 2.6) and then, in §4.2, define the preferred principal value selected by the local trace formula.

We let K be a local field. The natural representation $\vartheta_{\rm a}$ of K^* on $L^2(K)$ is of the form (2.51) i.e.

$$(\vartheta_{\rm a}(\lambda)\,\xi)(q) = \xi(\lambda^{-1}q)\,, \tag{2.71}$$

and to a test function h one associates the operator

$$\vartheta_{\rm a}(h) = \int_{K^*} h(\lambda)\vartheta_{\rm a}(\lambda)\,d^*\lambda \tag{2.72}$$

where $d^*\lambda$ is the Haar measure on the multiplicative group K^*. The latter has a preferred normalization which will be discussed in §4.1 below.
We first give a heuristic explanation for the integrand $\frac{h(u^{-1})}{|1-u|}\,d^*u$ in (2.70). We explain why at the formal level the distribution-theoretic trace of $\vartheta_{\rm a}(\lambda^{-1})$ is given by $\frac{1}{|1-\lambda|}$. This only makes sense when $\lambda \neq 1$ and the precise justification requires the discussion of functoriality and transversality properties of distributions (cf. [**156**] and §7.1 of Chapter 4). The formal computation of the distribution-theoretic trace of $\vartheta_{\rm a}(\lambda^{-1})$ is given by the integral on the diagonal $\int k(x,x)dx$ of the associated Schwartz kernel, which is $k(x,y)dy = \delta(y-\lambda\,x)\,dy$ since

$$(\vartheta_{\rm a}(\lambda^{-1})\,\xi)(x) = \xi(\lambda x) = \int \delta(y-\lambda\,x)\,\xi(y)\,dy = \int k(x,y)\,\xi(y)\,dy\,. \tag{2.73}$$

Thus this gives

$$\int k(x,x)\,dx = \int \delta(x-\lambda\,x)\,dx = \int \delta((1-\lambda)x)\,dx$$
$$= \int \delta(y)\,d((1-\lambda)^{-1}\,y) = |1-\lambda|^{-1}\int \delta(y)\,dy = |1-\lambda|^{-1}\,.$$

One can reformulate this result as asserting that the product of the distributions $\delta(x-y)$ and $\delta(y-\lambda\,x)$ gives $|1-\lambda|^{-1}\,\delta_0$ so that

$$\int \delta(x-y)\,\delta(y-\lambda\,x)\,dx\,dy = |1-\lambda|^{-1}\,. \tag{2.74}$$

The product of distributions u and v only makes sense under the hypothesis $0 \notin \mathrm{WF}(u)+\mathrm{WF}(v)$ where $\mathrm{WF}(u)$ is the wave front set of u (cf. [**156**] Proposition 3.10). In the above context of local fields, the correct way of defining the product of two distributions is as the inverse Fourier transform of the convolution of the Fourier transforms. We shall explain how to perform the computation for the product of $\delta(x-y)$ and $\delta(y-\lambda\,x)$ in order to understand the condition $\lambda \neq 1$ as a transversality condition $0 \notin \mathrm{WF}(u)+\mathrm{WF}(v)$. We let α be a nontrivial character of the additive group K and use the pairing between K^2 and its dual K^2 given by the bicharacter

$$\tilde\alpha((x,y),(\xi,\eta)) = \alpha(x\,\xi + y\,\eta) \in U(1)\,. \tag{2.75}$$

LEMMA 2.7. *Let $(a,b) \in K^2(\neq 0)$. The Fourier transform $\mathbf{F}_{\tilde\alpha}$ of $\delta(ax+by)$ is $\delta(-b\,\xi + a\,\eta)$.*

PROOF. One first computes the Fourier transform of $\delta(x)$; the additive Haar measure dx is normalized so as to be selfdual, and in one variable, $\delta(x)$ and 1 are Fourier transforms of each other. Thus, with $\mathbf{F}_{\tilde\alpha}(f) = f^\wedge$ the Fourier transform, one gets

$$(\delta\otimes 1)^\wedge = 1\otimes\delta\,. \tag{2.76}$$

Let then $(c,d) \in K^2$ be such that $ad - bc = 1$ and consider the invertible linear transformation of K^2,

$$L \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{2.77}$$

The Fourier transform $\mathbf{F}_{\tilde{\alpha}}(f) = f^\wedge$ obeys the rule

$$(f \circ L)^\wedge = |\det L|^{-1} f^\wedge \circ (L^{-1})^t . \tag{2.78}$$

Here one has $\det L = 1$ and $(L^{-1})^t$ is

$$(L^{-1})^t = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}. \tag{2.79}$$

Thus (2.78) gives the Fourier transform of $\delta(ax+by)$ as $\delta(-b\,\xi + a\,\eta)$. ☐

Thus, to justify (2.74), we have to compute the convolution of the two generalized functions $\delta(\xi+\eta)$ and $\delta(\xi+\lambda\,\eta)$. Now

$$\int f(\xi,\eta)\,\delta(\xi+\eta)\,d\xi\,d\eta = \int f(\xi,-\xi)\,d\xi \tag{2.80}$$

and

$$\int f(\xi,\eta)\,\delta(\xi+\lambda\,\eta)\,d\xi\,d\eta = \int f(-\lambda\,\eta,\eta)\,d\eta\,. \tag{2.81}$$

Thus we are dealing with two measures carried, respectively, by the two distinct lines W_1 and W_λ where $W_\alpha = \{(-\alpha\,\eta,\eta)\,|\,\eta\in K\}$, $\forall \alpha \in K$. The transversality condition $0 \notin \mathrm{WF}(u) + \mathrm{WF}(v)$ is here $0 \notin W_1^* + W_\lambda^*$, where $W_\alpha^* = W_\alpha \setminus \{0\}$. It holds if and only if $\lambda \neq 1$. The convolution of these measures evaluated on $f \in C_c^\infty(K^2)$ is

$$\int f(\alpha+\beta)\,d\mu(\alpha)\,d\nu(\beta) = \int\int f((\xi,-\xi)+(-\lambda\,\eta,\eta))\,d\xi\,d\eta$$

$$= \int\int f(\xi-\lambda\,\eta,-\xi+\eta)\,d\xi\,d\eta = \Big(\int\int f(\xi',\eta')d\xi'\,d\eta'\Big)\times|J|^{-1}$$

where J is the determinant of the matrix $\begin{bmatrix} 1 & -\lambda \\ -1 & 1 \end{bmatrix} = L$, so that $\begin{bmatrix} \xi' \\ \eta' \end{bmatrix} = L \begin{bmatrix} \xi \\ \eta \end{bmatrix}$. One has $J = 1-\lambda$ and thus the convolution of the generalized functions $\delta(\xi+\eta)$ and $\delta(\xi+\lambda\,\eta)$ gives as expected the constant function

$$|1-\lambda|^{-1}\,1\,. \tag{2.82}$$

4.1. Normalization of Haar measure on a modulated group.

In order to fix the notation we first recall the following:

Definition 2.8.

(1) *A modulated group G is a locally compact group with a proper morphism with cocompact range,*

$$\mathrm{Mod}\ :\ g \to |g|\,,\ G \to \mathbb{R}_+^*\,. \tag{2.83}$$

(2) *A Haar measure d^*g on G is normalized when*

$$\int_{|g|\in[1,\Lambda]} d^*g \sim \log\Lambda \qquad \textit{when } \Lambda \to +\infty\,. \tag{2.84}$$

We let ρ_Λ be the function on G given by

$$\rho_\Lambda(g) = 0 \quad \text{if } |g| \notin [1,\Lambda]\,, \;\; \rho_\Lambda(g) = \frac{1}{\log \Lambda} \quad \text{if } g \in [1,\Lambda]\,. \tag{2.85}$$

The normalization (2.84) means that $\int \rho_\Lambda \, d^*g \to 1$ when $\Lambda \to \infty$. There exists a unique normalized Haar measure d^*g on G. Let

$$G_0 = \text{Ker Mod} = \{g \in G \,|\, |g| = 1\}\,.$$

It is a compact group by hypothesis, and one can identify G/G_0 with the range N of the module. Let us determine the measure d^*n on $N \subset \mathbb{R}_+^*$ such that (2.84) holds for

$$\int f(g)\, d^*g = \int \left(\int f(ng_0)\, dg_0 \right) d^*n \tag{2.86}$$

where the Haar measure dg_0 on G_0 is normalized by

$$\int_{G_0} dg_0 = 1\,, \tag{2.87}$$

and the function $\int f(ng_0)\, dg_0$ only depends on $|n| \in N$.

Proposition 2.9. *For $N = \mathbb{R}_+^*$ the normalized Haar measure is*

$$d^*\lambda = \frac{d\lambda}{\lambda}\,. \tag{2.88}$$

For $N = \mu^{\mathbb{Z}}$, for some $\mu > 1$, the normalized Haar measure is

$$\int f(g)\, d^*g = \log \mu \sum_{n \in \mathbb{Z}} f(\mu^n)\,. \tag{2.89}$$

Proof. The first statement is clear. For $f = \rho_\Lambda$ the right hand side of (2.89) is $\log \mu \, \frac{k}{\log \Lambda}$ where k is the number of $\mu^n \in [1,\Lambda]$, thus $k \sim \frac{\log \Lambda}{\log \mu}$. This shows that (2.84) holds. ☐

Proposition 2.10. *Let $H \subset G$ be a cocompact subgroup of an abelian modulated group G. Then H is modulated by the restriction of* Mod *to H. Let both d^*g and d^*h be normalized by* (2.84)*; then one has*

$$\int \left(\int f(hy) d^*h \right) d_0 y = \int f d^*g \tag{2.90}$$

where $d_0 y$ is the Haar measure of integral 1 on G/H,

$$\int_{G/H} d_0 y = 1\,. \tag{2.91}$$

Proof. The left-hand side of (2.90) defines a Haar measure on G and we just need to show that it satisfies (2.84). Let ρ_Λ^G be defined by (2.85) for G. For any $y \in G$ one has $|yh| = |y||h|$ and $\rho_\Lambda^G(hy)$ agrees with $\rho_\Lambda^H(h)$ except near the boundary, so that

$$\int |\rho_\Lambda^G(hy) - \rho_\Lambda^H(h)| d^*h \to 0 \text{ when } \Lambda \to \infty$$

and

$$\int \rho_\Lambda(hy)\, d^*h \to 1 \qquad \text{when } \Lambda \to \infty \tag{2.92}$$

uniformly on compact sets of $y \in G$. Thus

$$\int \left(\int \rho_\Lambda(hy)\, d^*h \right) d_0 y \to 1 \qquad \text{when } \Lambda \to \infty \tag{2.93}$$

which shows that the left-hand side of (2.90) defines a normalized Haar measure on G. □

4.2. Principal values.

We now describe the principal value dictated by the local trace formula (2.70). We let K be a local field. The multiplicative group K^* is a modulated group for the module given by

$$\mathrm{Mod} \; : \; x \to |x| \,, \;\; d(xa) = |x| da \,, \tag{2.94}$$

for any Haar measure da on the additive group of K. A *basic character* α of K is, by definition, a non-trivial character of the additive group K. Given such an α the associated Fourier transform is

$$(\mathbf{F}_\alpha \rho)(y) = \hat{\varrho}(y) = \int \varrho(x) \alpha(xy)\, dx \,. \tag{2.95}$$

which makes sense for any element $f \in \mathcal{S}(K)$ of the Bruhat-Schwartz space $\mathcal{S}(K)$ (cf. [**36**] for the definition of $\mathcal{S}(K)$). By the Plancherel Theorem $\mathbf{F}_\alpha$ extends to a unitary operator

$$\langle \mathbf{F}_\alpha(\xi), \mathbf{F}_\alpha(\eta) \rangle = \langle \xi, \eta \rangle \,, \quad \forall \xi, \eta \in L^2(K) \,, \tag{2.96}$$

where one has normalized the additive Haar measure da on K so that it is self-dual. The adjoint of the operator $\mathbf{F}_\alpha$ is $\mathbf{F}_{\bar{\alpha}}$, where $\bar{\alpha}(a) = \alpha(-a)$, and it follows from the Parseval formula (2.96) that $\mathbf{F}_{\bar{\alpha}}$ is the inverse of $\mathbf{F}_\alpha$. The transpose of the the operator $\mathbf{F}_\alpha$ is $\mathbf{F}_\alpha$ and this gives its natural extension to tempered distributions, i.e. continuous linear forms on the Bruhat-Schwartz space $\mathcal{S}(K)$,

$$(\mathbf{F}_\alpha(L))(f) := L(\mathbf{F}_\alpha(f)) \,, \quad \forall f \in \mathcal{S}(K) \,, \;\; L \in \mathcal{S}'(K) \,. \tag{2.97}$$

PROPOSITION 2.11. *Let K be a local field and α a basic character.*

(1) *Let da be the selfdual Haar measure on K. The Fourier transform $\mathbf{F}_\alpha(\varphi) = \hat{\varphi}$ of the distribution $\varphi(u) = -\log|u|$ is given away from 0 by*

$$\int f(a)\, \hat{\varphi}(a)\, da = \int f(a)\, d^*a \,, \;\; \forall f \,, \quad f(0) = 0 \,, \tag{2.98}$$

where d^*a is the normalized Haar measure on K^*.

(2) *There exists a unique distribution ϱ_α extending d^*u at $u = 0$, whose Fourier transform satisfies the vanishing condition $\mathbf{F}_\alpha(\varrho_\alpha)(1) = 0$.*

(3) *One has, for $\lambda \in K^*$, $f \in \mathcal{S}(K)$, $f_\lambda(x) = f(\lambda x)$,*

$$\varrho_\alpha(\alpha) = 0 \,, \quad \varrho_\alpha(f_\lambda) = \varrho_\alpha(f) - \log|\lambda|\, f(0) \,. \tag{2.99}$$

(4) *Let $a \in K^*$ and β be such that $\beta(x) = \alpha(ax)$ for all $x \in K$; then*

$$\varrho_\beta = \varrho_\alpha + \log|a|\, \delta_0 \,. \tag{2.100}$$

Proof. (1) To see this one lets P be the distribution on K given by

$$P(f) = \lim_{\varepsilon \in \mathrm{Mod}(K), \varepsilon \to 0} \left(\int_{|x| \geq \varepsilon} f(x)\, d^*x + f(0) \log \varepsilon \right). \tag{2.101}$$

The normalization of the measure d^*a ensures that the limit exists when f is locally constant near 0 so that it makes sense for any Schwartz function $f \in \mathcal{S}(K)$. For $\lambda \in K^*$ let $f_\lambda(x) = f(\lambda x)$. One has $P(f_\lambda) = P(f) - \log|\lambda|\, f(0)$. One has

$$P(f_\lambda) = \int \hat{P}(x)\, (\widehat{f_\lambda})(x) dx = \int \hat{P}(\lambda\, y) \hat{f}(y)\, dy$$

since the Fourier transform of f_λ is $\frac{1}{|\lambda|} \hat{f}_{\lambda^{-1}}$. This is enough to show that the function $h = \hat{P}(x)$ fulfills $h(\lambda\, x) = h(x) - \log|\lambda|$, and is equal to $-\log|x|$ up to an additive constant. More precisely, both $\hat{P}(x)$ and $-\log|x|$ are tempered distributions and their difference fulfills the same homogeneity condition as the constants, and is thus a constant by [**296**]. Thus φ differs from $\hat{P}$ by a constant and $\hat{\varphi}$ differs from P by a multiple of δ_0.
(2) The equality

$$f \mapsto \int f(u) d^*u\,, \quad \forall f \in \mathcal{S}(K)\,, f(0) = 0$$

defines a continuous linear form on the hyperplane $\{f | f(0) = 0\}$ of the Bruhat-Schwartz space $\mathcal{S}(K)$. By the Hahn-Banach theorem this extends to a continuous linear form on $\mathcal{S}(K)$ and hence a distribution L. The first part of the proposition gives the existence of ρ_α. Two extensions L differ by a scalar multiple of the Dirac mass $f \mapsto f(0)$. The Fourier transform of the Dirac mass is the constant function and thus one gets the uniqueness of ϱ_α.
(3) This follows from the proof of (1), using $P(f_\lambda) = P(f) - \log|\lambda|\, f(0)$.
(4) The difference $\varrho_\beta - \varrho_\alpha$ is a scalar multiple of δ_0. To show that this scalar is $\log|a|$ we apply both sides to $\beta = \alpha_a$ and get, using (2.99),

$$(\varrho_\beta - \varrho_\alpha)(\beta) = \varrho_\beta(\beta) - \varrho_\alpha(\alpha_a) = -\varrho_\alpha(\alpha_a) = \log|a|\,.$$

∎

Definition 2.12. *The principal value $\int'$ is defined by*

$$\int'_{(K,\alpha)} \frac{f(u^{-1})}{|1-u|}\, d^*u = \langle \varrho_\alpha, g \rangle, \qquad \textit{with} \quad g(\lambda) = \frac{f((\lambda+1)^{-1})}{|\lambda+1|}, \tag{2.102}$$

where $\langle \varrho_\alpha, g \rangle$ is the pairing of the distribution ϱ_α and the function $g(\lambda)$.

This makes sense provided the support of f is compact, which implies that $g(\lambda)$ vanishes identically in a neighborhood of $\lambda = -1$.

Lemma 2.13. *The principal value $\int'$ is given, after a translation of the singularity to $a = 0$, by*

$$\frac{1}{\rho} \int' g(a)\, \frac{da}{|a|} = -\int \hat{g}(u)\, \log|u|\, du\,, \quad d^*a = \frac{da}{\rho\,|a|}\,. \tag{2.103}$$

Proof. The normalization condition $\hat{L}(1) = 0$ on a distribution L on K is invariant under translations of the variable since this multiplies the Fourier transform $\hat{L}$ by a character and the result still vanishes at $1 \in K$. Thus since $\log|u|$ vanishes at $1 \in K$ we get (2.103). ∎

We now consider more explicitly the case $K = \mathbb{R}$ and we choose, as above, the additive character $e_{\mathbb{R}}(x) = e^{-2\pi i x}$ so that the Fourier transform is

$$(\mathbf{F}_{e_{\mathbb{R}}} f)(x) = \int f(y)\, e^{-2\pi i x y}\, dy \tag{2.104}$$

and the Haar measure dx is selfdual. The normalized Haar measure on $K^* = \mathbb{R}^*$ is

$$d^*x = \frac{1}{2}\frac{dx}{|x|} \ . \tag{2.105}$$

PROPOSITION 2.14.

(1) *The Fourier transform of* $-\log|p|$ *is given by the distribution*

$$L(f) := (\log 2\pi + \gamma) f(0) + \lim_{\varepsilon \to 0}\left(\int_{|x|>\varepsilon} f(x) d^*x + (\log \varepsilon) f(0)\right) , \tag{2.106}$$

where γ *is the Euler constant.*

(2) *Let* $\lambda = \log 2\pi + \gamma$*; then one has for any test function* f,

$$\begin{aligned} \textstyle\int'_{\mathbb{R}} f(u)\, \frac{1}{|1-u|}\, d^*u &= \lambda\, f(1) + \lim_{\varepsilon\to 0}\left(\int \frac{f(u)}{|1-u|^{1-\varepsilon}}\, d^*u - \frac{1}{\varepsilon}\, f(1)\right) \\ &= \lambda\, f(1) + \lim_{\varepsilon\to 0}\left(\int_{|1-u|\geq\varepsilon} \frac{f(u)}{|1-u|}\, d^*u + (\log\varepsilon)\, f(1)\right) . \end{aligned} \tag{2.107}$$

PROOF. (1) Let $\varphi(u) = -\log|u|$. It is a tempered distribution on $\mathbb{R}$ and

$$\langle \varphi, e^{-\pi u^2}\rangle = \int (-\log|u|) e^{-\pi u^2} dy = \frac{1}{2}\log\pi + \frac{\gamma}{2} + \log 2 \tag{2.108}$$

as one obtains from $\frac{\partial}{\partial s}\int |u|^{-s}\, e^{-\pi u^2}\, du = \frac{\partial}{\partial s}\left(\pi^{\frac{s-1}{2}}\,\Gamma\left(\frac{1-s}{2}\right)\right)$ evaluated at $s = 0$, using $\frac{\Gamma'(\frac{1}{2})}{\Gamma(\frac{1}{2})} = -\gamma - 2\log 2$. Thus by the Parseval formula (2.96),

$$\langle \hat\varphi, e^{-\pi x^2}\rangle = \frac{1}{2}\log\pi + \frac{\gamma}{2} + \log 2\, . \tag{2.109}$$

The distributions $\hat\varphi$, L and the following

$$\mathrm{MS}(f) = \lim_{\varepsilon\to 0}\left(\int f(x)\, |x|^{\varepsilon}\, d^*x - \frac{1}{\varepsilon}\, f(0)\right) \tag{2.110}$$

agree on functions $f \in \mathcal{S}(\mathbb{R})$ which vanish at $0 \in \mathbb{R}$ since $f\, d^*x$ is then integrable and Lemma 2.13 applies. Moreover one has

$$\begin{aligned} &\lim_{\varepsilon\to 0}\left(\int_{|x|\geq\varepsilon} f(x)\, d^*x + (\log\varepsilon)\, f(0)\right) \\ &= \lim_{\varepsilon\to 0}\left(\int f(x)\, |x|^{\varepsilon}\, d^*x - \frac{1}{\varepsilon}\, f(0)\right) . \end{aligned} \tag{2.111}$$

Indeed the difference between these two distributions is a multiple of the Dirac mass δ_0 while both sides make sense and vanish for $f(x) = 1$ if $|x| \leq 1$, $f(x) = 0$ otherwise, so that they agree on any test function extending f smoothly. One has

$$\hat\varphi = \mathrm{MS} + \lambda\, \delta_0\, , \quad \lambda = \log 2\pi + \gamma\, . \tag{2.112}$$

Indeed for $f(x) = e^{-\pi x^2}$ one has

$$\int f(x)\, |x|^{\varepsilon}\, d^*x = \frac{1}{2}\int_{\mathbb{R}} e^{-\pi x^2}\frac{dx}{|x|^{1-\varepsilon}} = \frac{1}{2}\pi^{-\varepsilon/2}\Gamma(\frac{\varepsilon}{2}) = \frac{1}{\varepsilon} - \frac{\gamma}{2} - \frac{\log\pi}{2} + O(\varepsilon)$$

which gives the same answer as (2.109) after adding $\lambda f(0)$. Using (2.111) we thus get that, for any test function f,

$$\langle \hat{\varphi}, f \rangle = \lim_{\varepsilon \to 0} \left(\int_{|x| \geq \varepsilon} f(x)\, d^*x + (\log \varepsilon)\, f(0) \right) + \lambda\, f(0) \tag{2.113}$$

where $\lambda = \log(2\pi) + \gamma$.

(2) From (2.113) and (2.111) one gets (2.107). □

COROLLARY 2.15. *The principal value $\int'$ is given explicitly for functions f on $\mathbb{R}^*$ fulfilling $f(-u) = f(u) = f(u^{-1})$ by the formula*

$$\begin{gathered} \textstyle\int' f(u) \frac{|u|^{1/2}}{|1-u|} d^*u = (\log \pi + \gamma) f(1) \\ + \lim_{\varepsilon \to 0} \left(\textstyle\int_0^1 f(u) \frac{2u^{1/2}}{(1-u^2)^{1-\varepsilon}} \frac{du}{u} - \frac{1}{\varepsilon} f(1) \right) \end{gathered} \tag{2.114}$$

where γ is the Euler constant.

PROOF. To see why (2.114) holds, we use Proposition 2.14 which gives

$$\begin{gathered} \textstyle\int' f(u) \frac{|u|^{1/2}}{|1-u|} \frac{du}{2|u|} = (\log 2\pi + \gamma) f(1) \\ + \lim_{\varepsilon \to 0} \left(\textstyle\int_{|1-u| > \varepsilon} f(u) \frac{|u|^{1/2}}{|1-u|} \frac{du}{2|u|} + (\log \varepsilon) f(1) \right) . \end{gathered} \tag{2.115}$$

Since we have $f(u) = f(u^{-1})$, the second term in (2.115) equals

$$2 \lim_{\varepsilon \to 0} \left(\int_{|u|<1,\, |1-u|>\varepsilon} f(u) \frac{|u|^{1/2}}{|1-u|} \frac{du}{2|u|} + \frac{1}{2} (\log \varepsilon) f(1) \right) . \tag{2.116}$$

By hypothesis we have $f(-u) = f(u)$, so that we can write (2.116) as

$$\lim_{\varepsilon \to 0} \left(\int_0^{1-\varepsilon} f(u) \frac{2u^{1/2}}{(1-u^2)} \frac{du}{u} + (\log \varepsilon) f(1) \right) . \tag{2.117}$$

Finally, we can check, using $f(u) = u^{3/2}$ for $u \in [0,1]$, that one can rewrite this as

$$\lim_{\varepsilon \to 0} \left(\int_0^1 f(u) \frac{2u^{1/2}}{(1-u^2)^{1-\varepsilon}} \frac{du}{u} - \frac{1}{\varepsilon} f(1) \right) - f(1) \log 2. \tag{2.118}$$

This proves Corollary 2.15. □

We can also express this finite value using the following variant of the minimal subtraction scheme:

COROLLARY 2.16. *One has with $\mu = \frac{e^{-\gamma}}{2\pi}$,*

$$\int_{\mathbb{R}}' f(u) \frac{1}{|1-u|} d^*u = \lim_{\varepsilon \to 0} \left(\int \frac{f(u)}{|1-u|^{1-\varepsilon}} \mu^{-\varepsilon} d^*u - \frac{1}{\varepsilon} f(1) \right) . \tag{2.119}$$

In other words the principal value is obtained by a simple subtraction of the pole part as in dimensional regularization, using a scheme similar to $\overline{\mathrm{MS}}$, which would use $\mu^2 = \frac{e^{\gamma}}{4\pi}$.

5. Quantum states of the scaling flow

We now return to the problem of counting the number of states of the quantum mechanical system given by scaling transformations on the Hilbert space $\mathcal{H} = L^2(\mathbb{R})^{\text{even}}$. By construction the function h_E of (2.67) is a function on the multiplicative group $\mathbb{R}^*$. This group is a modulated group (Definition 2.8) with module given by $u \mapsto |u|$ and its normalized Haar measure (2.84) is given by

$$d^*\lambda := \frac{1}{2}\frac{d\lambda}{|\lambda|}\,. \tag{2.120}$$

In Theorem 2.6 the Hilbert space is $L^2(\mathbb{R})$, but we apply ϑ_{a} to functions h fulfilling the condition

$$h(-u) = h(u)\,, \quad \forall u \in \mathbb{R}^*.$$

The range of $\vartheta_{\mathrm{a}}(h)$ is then contained in $L^2(\mathbb{R})^{\text{even}}$. Thus, the formula (2.70) gives the trace in our quantum mechanical Hilbert space $\mathcal{H} = L^2(\mathbb{R})^{\text{even}}$.
If we proceed to apply (2.70) to the function h_E of (2.67) (note that we still need to justify this point since h_E does not have compact support) we find that the first term gives

$$2h_E(1)\log\Lambda = \frac{1}{2\pi} \times 2E \times 2\log\Lambda. \tag{2.121}$$

We then need to compute precisely the second term

$$\int' \frac{h_E(u^{-1})}{|1-u|}\, d^*u. \tag{2.122}$$

We obtain the following result.

Proposition 2.17. *The second term of* (2.70) *applied to the function* $h = h_E$ *of* (2.67) *is*

$$\int' \frac{h_E(u^{-1})}{|1-u|}\, d^*u = -2(\langle N(E)\rangle - 1). \tag{2.123}$$

Proof. By definition we have

$$\langle N(E)\rangle = N(E) - N_{\text{osc}}(E), \tag{2.124}$$

where

$$N_{\text{osc}}(E) = \frac{1}{\pi}\Im \log\zeta\left(\frac{1}{2} + iE\right). \tag{2.125}$$

To complete the proof of Proposition 2.17, we apply (2.114) to

$$f_s(u) = |u|^{is} + |u|^{-is}. \tag{2.126}$$

We need to compute

$$\int_0^1 u^{is}\frac{2u^{1/2}}{(1-u^2)^{1-\varepsilon}}\frac{du}{u}\,, \tag{2.127}$$

With $x = u^2$, we get

$$\int_0^1 x^{\left(\frac{1}{4}+i\frac{s}{2}\right)}(1-x)^{-1+\varepsilon}\frac{dx}{x} = B\left(\frac{1}{4}+i\frac{s}{2},\varepsilon\right) = \frac{\Gamma\left(\frac{1}{4}+i\frac{s}{2}\right)\Gamma(\varepsilon)}{\Gamma\left(\frac{1}{4}+i\frac{s}{2}+\varepsilon\right)}\,. \tag{2.128}$$

The residue at $\varepsilon = 0$ is equal to 1 and the finite part gives

$$-\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}+i\frac{s}{2}\right)-\gamma\,. \tag{2.129}$$

Thus, we obtain

$$\begin{array}{c} \int' f_s(u)\frac{|u|^{1/2}}{|1-u|}d^*u \\ = 2(\log\pi+\gamma)-2\gamma-\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}+i\frac{s}{2}\right)-\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}-i\frac{s}{2}\right)\,. \end{array} \tag{2.130}$$

Finally, this gives

$$\int' \frac{h_E(u^{-1})}{|1-u|}d^*u = \frac{1}{2\pi}\int_0^E\left(2\log\pi-\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}+i\frac{s}{2}\right)-\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}-i\frac{s}{2}\right)\right)ds\,.$$

Using (2.31), we then get

$$\int' \frac{h_E(u^{-1})}{|1-u|}d^*u = -2(\langle N(E)\rangle - 1)\,. \tag{2.131}$$

□

Thus, we find the following asymptotic formula for the number of quantum states when $\Lambda \to \infty$.

THEOREM 2.18. *The dimension of the near intersection of Q_Λ with N_E is given by*

$$\mathrm{Tr}(Q_\Lambda N_E) = \frac{4E}{2\pi}\log\Lambda - 2(\langle N(E)\rangle - 1) + o(1), \quad \textit{for}\ \ \Lambda\to\infty. \tag{2.132}$$

This still requires justifying the fact that we applied the result of Theorem 2.6 to the function $h_E \notin \mathcal{S}(\mathbb{R}^*)$. This is discussed in §5.1 below. We also give in Remark 2.23 below an explanation for the additive 2 that appears in the expression $-2\langle N(E)\rangle+2$ in (2.131) and (2.132).

5.1. Quantized calculus.

In order to prove Theorem 2.18 and refine the analysis of §5, we use the quantized calculus developed in [**68**]. In particular, we analyze here the relative position of the three projections P_Λ, $\widehat{P}_\Lambda$, and N_E, using identities involving the quantized calculus, as proved in [**72**]. The method used here is based on the idea of Burnol [**37**] which simplifies the original argument of [**71**].

First recall from §IV of [**68**] that the main idea of quantized calculus is to give an operator-theoretic version of the calculus rules, based on the operator-theoretic differential

$$đf := [F, f], \tag{2.133}$$

where f is an element in an involutive algebra $\mathcal{A}$ represented as bounded operators on some Hilbert space $\mathcal{H}$, and the right-hand side of (2.133) is the commutator with a self-adjoint operator F on $\mathcal{H}$ with $F^2 = 1$.

In particular, we recall the framework for the quantized calculus in one variable, as in §IV of [**68**]. We let functions $f(s)$ of one real variable s act as multiplication operators on $L^2(\mathbb{R})$, by

$$(f\,h)(s) := f(s)\,h(s)\,,\quad \forall h\in L^2(\mathbb{R}). \tag{2.134}$$

We let $\mathbf{F}_{e_\mathbb{R}}$ denote the Fourier transform with respect to the basic character $e_\mathbb{R}(x) = e^{-2\pi i x}$, namely

$$\mathbf{F}_{e_\mathbb{R}}(h)(y) := \int h(x)\, e^{-2\pi i x y}\, dy. \tag{2.135}$$

We also introduce the notation

$$\Pi_{[a,b]} := \mathbf{F}_{e_\mathbb{R}}\, \mathbf{1}_{[a,b]} \mathbf{F}_{e_\mathbb{R}}^{-1}, \tag{2.136}$$

for the conjugate by the Fourier transform $\mathbf{F}_{e_\mathbb{R}}$ of the multiplication operator by the characteristic function $\mathbf{1}_{[a,b]}$ of the interval $[a,b] \subset \mathbb{R}$.

Definition 2.19. *We define the* quantized differential *of f to be the operator*

$$đf := [H,\, f] = H\, f - f\, H, \tag{2.137}$$

where H is the Hilbert transform $H = 2\,\Pi_{[0,\infty]} - 1$ given by

$$(H\, h)(s) := \frac{1}{i\pi} \int \frac{h(t)}{s-t} dt. \tag{2.138}$$

Thus, the quantized differential of f is given by the kernel

$$k(s,t) = \frac{i}{\pi}\, \frac{f(s) - f(t)}{s-t}. \tag{2.139}$$

We follow [**281**] and [**37**], and use the classical formula expressing the Fourier transform as a composition of the inversion

$$I(f)(s) := f(s^{-1}) \tag{2.140}$$

with a multiplicative convolution operator. We use the unitary identification

$$w : L^2(\mathbb{R}, ds)^{\text{even}} \to L^2(\mathbb{R}_+^*, d^*\lambda), \;\; w(\eta)(\lambda) := \lambda^{1/2}\, \eta(\lambda), \forall \lambda \in \mathbb{R}_+^* \tag{2.141}$$

whose inverse is given by

$$w^{-1} : L^2(\mathbb{R}_+^*, d^*\lambda) \to L^2(\mathbb{R}, ds)^{\text{even}}, \;\; w^{-1}(\xi)(x) := |x|^{-1/2}\, \xi(|x|)\,. \tag{2.142}$$

Also we define the duality $\langle \mathbb{R}_+^*, \mathbb{R} \rangle$ by the bicharacter

$$\mu(v,s) = v^{-is}\,, \quad \forall v \in \mathbb{R}_+^*, s \in \mathbb{R}\,, \tag{2.143}$$

so that the Fourier transform $\mathbf{F}_\mu : L^2(\mathbb{R}_+^*) \to L^2(\mathbb{R})$ associated to the bicharacter μ is

$$\mathbf{F}_\mu(f)(s) := \int_0^\infty f(v) v^{-is} d^*v\,. \tag{2.144}$$

Lemma 2.20. *On $L^2(\mathbb{R})^{\text{even}}$ one has*

$$\mathbf{F}_{e_\mathbb{R}} = w^{-1} \circ I \circ \mathbf{F}_\mu^{-1} \circ u \circ \mathbf{F}_\mu \circ w, \tag{2.145}$$

where u is the multiplication operator by the function

$$u(s) := e^{2\, i\, \theta(s)}, \tag{2.146}$$

where $\theta(s)$ is the Riemann-Siegel angular function of (2.24).

PROOF. First $\mathbf{F}_{e_\mathbb{R}}$ preserves globally $L^2(\mathbb{R})^{\rm even}$. One has, for $\xi \in L^2(\mathbb{R}_+^*)$,

$$\begin{aligned}(w \circ \mathbf{F}_{e_\mathbb{R}} \circ w^{-1})(\xi)(v) &= v^{1/2} \int_\mathbb{R} |x|^{-1/2}\, \xi(|x|)\, e^{-2\pi i x v}\, dx \\ &= v^{1/2} \int_{\mathbb{R}_+^*} u^{1/2}\, \xi(u)\, (e^{2\pi i u v} + e^{-2\pi i u v})\, d^* u\,.\end{aligned}$$

This gives

$$\begin{aligned}(I \circ w \circ \mathbf{F}_{e_\mathbb{R}} \circ w^{-1})(\xi)(\lambda) &= (w \circ \mathbf{F}_{e_\mathbb{R}} \circ w^{-1})(\xi)(\lambda^{-1}) \\ &= \lambda^{-1/2} \int_{\mathbb{R}_+^*} (e^{2i\pi u \lambda^{-1}} + e^{-2i\pi u\lambda^{-1}})\, u^{1/2} \xi(u) d^* u\end{aligned}$$

We thus obtain

$$I \circ w \circ \mathbf{F}_{e_\mathbb{R}} \circ w^{-1} = C \tag{2.147}$$

where the operator C on the Hilbert space $L^2(\mathbb{R}_+^*, d^*\lambda)$ is given by convolution by

$$v \mapsto 2v^{-1/2} \cos(2\pi v^{-1}). \tag{2.148}$$

By construction C is unitary and commutes with the regular representation of $\mathbb{R}_+^*$. Thus $C = \mathbf{F}_\mu^{-1} \circ u \circ \mathbf{F}_\mu$ where u is the Fourier transform

$$u(s) = \int_0^\infty 2v^{-1/2} \cos(2\pi v^{-1}) v^{-is} d^* v. \tag{2.149}$$

This is defined in the sense of distributions. To compute (2.149) one can let $v = e^{-t}$ and use

$$\int_\mathbb{R} e^{t/2} e^{-z e^t} e^{ist}\, dt = \, z^{-(1/2+is)} \Gamma(1/2+is)\,, \quad \forall z\,,\ \Im z > 0\,. \tag{2.150}$$

This gives

$$u(s) = \, 2\cos((1/2+is)\pi/2)(2\pi)^{-(1/2+is)}\Gamma(1/2+is) \tag{2.151}$$

and the duplication formula

$$\Gamma\left(\frac{z}{2}\right)\Gamma\left(\frac{1+z}{2}\right) = \pi^{1/2} 2^{1-z} \Gamma(z) \tag{2.152}$$

shows that $u(s)$ is given by

$$u(s) = \frac{\pi^{-z/2}\Gamma(z/2)}{\pi^{-(1-z)/2}\Gamma((1-z)/2)}\,,\ z = 1/2 + is\,, \tag{2.153}$$

which, using (2.24) gives (2.146). □

The following lemma relates the quantized calculus to the analysis of the geometry of the three projections P_Λ, $\widehat{P}_\Lambda$ and N_E.

LEMMA 2.21. *For any Λ there is a unitary operator*

$$W = W_\Lambda : L^2(\mathbb{R})^{\rm even} \to L^2(\mathbb{R}), \tag{2.154}$$

such that, for any functions $h_j \in \mathcal{S}(\mathbb{R}_+^*)$, $j = 1, 2$, one has

$$\begin{aligned}&W \vartheta_{\rm a}(\tilde h_1) \widehat{P}_\Lambda P_\Lambda \vartheta_{\rm a}(\tilde h_2) W^* = \\ &\hat h_1 \left(\tfrac{1}{2} u^{-1} đu \Pi_{[-\infty, \frac{2\log\Lambda}{2\pi}]} + \Pi_{[0, \frac{2\log\Lambda}{2\pi}]}\right) \hat h_2\end{aligned} \tag{2.155}$$

Here $\tilde h_j(\lambda)=\lambda^{-1/2}h_j(\lambda)$, the operator $\not{d}u$ is the quantized differential of the function u of (2.146) and $\hat h_j$ is the multiplication operator by the Fourier transform $\mathbf{F}_\mu(h_j)$.

PROOF. We let $\vartheta_{\rm m}$ be the regular representation of $\mathbb{R}_+^*$ on $L^2(\mathbb{R}_+^*)$

$$(\vartheta_{\rm m}(\lambda)\,\xi)(v) := \xi(\lambda^{-1}\,v)\,,\quad \forall \xi\in L^2(\mathbb{R}_+^*)\,. \tag{2.156}$$

One has

$$w\,\vartheta_{\rm a}(\tilde h_j)w^{-1}=\vartheta_{\rm m}(h_j),\quad \text{and}\quad \mathbf{F}_\mu\vartheta_{\rm m}(h_j)\mathbf{F}_\mu^{-1}=\widehat{h}_j\,. \tag{2.157}$$

Now we have $P_\Lambda=\mathbf{1}_{[-\Lambda,\Lambda]}$, so that $wP_\Lambda w^{-1}=\mathbf{1}_{[0,\Lambda]}$, and we obtain

$$\begin{aligned}\widehat{P}_\Lambda P_\Lambda=\mathbf{F}_{e_\mathbb{R}}P_\Lambda\mathbf{F}_{e_\mathbb{R}}^{-1}P_\Lambda &= w^{-1}I\mathbf{F}_\mu^{-1}u\mathbf{F}_\mu\mathbf{1}_{[0,\Lambda]}\mathbf{F}_\mu^{-1}u^{-1}\mathbf{F}_\mu I\mathbf{1}_{[0,\Lambda]}w\\ &= w^{-1}\mathbf{F}_\mu^{-1}u^{-1}\mathbf{F}_\mu\mathbf{1}_{[\Lambda^{-1},\infty]}\mathbf{F}_\mu^{-1}u\mathbf{F}_\mu\mathbf{1}_{[0,\Lambda]}w,\end{aligned} \tag{2.158}$$

where we used the identity $I\mathbf{F}_\mu^{-1}u\mathbf{F}_\mu=\mathbf{F}_\mu^{-1}u^{-1}\mathbf{F}_\mu I$, which follows from the symmetry $\theta(-s)=-\theta(s)$. We also used the identity $I\mathbf{1}_{[0,\Lambda]}I=\mathbf{1}_{[\Lambda^{-1},\infty]}$.
We now set

$$W_\Lambda := \mathbf{F}_\mu\,\vartheta_{\rm m}(\Lambda)\,w, \tag{2.159}$$

with $\vartheta_{\rm m}(\Lambda)$ as in (2.156). One has

$$W_\Lambda\,\vartheta_{\rm a}(\tilde h_1)\widehat{P}_\Lambda P_\Lambda\vartheta_{\rm a}(\tilde h_2)W_\Lambda^{-1}= \tag{2.160}$$

$$\begin{aligned}&\mathbf{F}_\mu\vartheta_{\rm m}(\Lambda)\vartheta_{\rm m}(h_1)\,\mathbf{F}_\mu^{-1}u^{-1}\mathbf{F}_\mu\mathbf{1}_{[\Lambda^{-1},\infty]}\mathbf{F}_\mu^{-1}u\mathbf{F}_\mu\mathbf{1}_{[0,\Lambda]}\vartheta_{\rm m}(h_2)\vartheta_{\rm m}(\Lambda)^{-1}\mathbf{F}_\mu^{-1}\\ &=\widehat{h}_1u^{-1}\mathbf{F}_\mu\vartheta_{\rm m}(\Lambda)\mathbf{1}_{[\Lambda^{-1},\infty]}\vartheta_{\rm m}(\Lambda)^{-1}\mathbf{F}_\mu^{-1}u\mathbf{F}_\mu\vartheta_{\rm m}(\Lambda)\mathbf{1}_{[0,\Lambda]}\vartheta_{\rm m}(\Lambda)^{-1}\mathbf{F}_\mu^{-1}\widehat{h}_2\,.\end{aligned}$$

Here we use the identities (2.157) and the fact that $\vartheta_{\rm m}(\Lambda)$ commutes with multiplicative convolution operators such as $\mathbf{F}_\mu^{-1}u^{-1}\mathbf{F}_\mu$. Next we see that

$$\vartheta_{\rm m}(\Lambda)\mathbf{1}_{[\Lambda^{-1},\infty]}\vartheta_{\rm m}(\Lambda)^{-1}=\mathbf{1}_{[1,\infty]}\,,\ \vartheta_{\rm m}(\Lambda)\mathbf{1}_{[0,\Lambda]}\vartheta_{\rm m}(\Lambda)^{-1}=\mathbf{1}_{[0,\Lambda^2]}\,. \tag{2.161}$$

In order to use the quantized calculus on functions on $\mathbb{R}$ as in Definition 2.19, we use the isomorphism of abelian groups

$$t\in\mathbb{R}\mapsto e^{2\pi t}\in\mathbb{R}_+^* \tag{2.162}$$

and note that $\mathbf{F}_\mu$ and $\mathbf{F}_{e_\mathbb{R}}$ are conjugate by this isomorphism since the bicharacter $\mu(v,s)=v^{-is}$ of (2.143) fulfills

$$\mu(e^{2\pi t},s)=e^{-2\pi ist}=e_\mathbb{R}(st)\,. \tag{2.163}$$

Thus we get from (2.136),

$$\mathbf{F}_\mu\mathbf{1}_{[a,b]}\mathbf{F}_\mu^{-1}=\Pi_{[\frac{\log a}{2\pi},\frac{\log b}{2\pi}]}\,, \tag{2.164}$$

and we obtain, using (2.160)

$$W_\Lambda\,\vartheta_{\rm a}(\tilde h_1)\widehat{P}_\Lambda P_\Lambda\vartheta_{\rm a}(\tilde h_2)W_\Lambda^{-1}=\widehat{h}_1u^{-1}\Pi_{[0,\infty]}\,u\,\Pi_{[-\infty,\frac{2\log\Lambda}{2\pi}]}\widehat{h}_2\,. \tag{2.165}$$

We then use

$$\frac{1}{2}\,\not{d}u=[\Pi_{[0,\infty]},u]\,, \tag{2.166}$$

which completes the proof of (2.155). ∎

5.2. Proof of Theorem 2.18.

As a first application of Lemma 2.21 we now complete the proof of Theorem 2.18. One has $N_E = \vartheta_{\rm m}(h_E)$, where $\widehat{h}_E = \mathbf{1}_{[-E,E]}$. Thus, $N_E \widehat{P}_\Lambda P_\Lambda$ is unitarily equivalent to

$$\mathbf{1}_{[-E,E]}\left(\frac{1}{2}u^{-1}\eth u\,\Pi_{[-\infty,\frac{2\log\Lambda}{2\pi}]} + \Pi_{[0,\frac{2\log\Lambda}{2\pi}]}\right). \tag{2.167}$$

The trace of $\mathbf{1}_{[-E,E]}\Pi_{[0,\frac{2\log\Lambda}{2\pi}]}$ is equal to $2E\frac{2\log\Lambda}{2\pi}$ and gives the leading term in the formula for $\mathrm{Tr}(N_E\widehat{P}_\Lambda P_\Lambda)$. If we replace $\Pi_{[-\infty,\frac{2\log\Lambda}{2\pi}]}$ by 1 the other term gives

$$\mathrm{Tr}\left(\mathbf{1}_{[-E,E]}\left(\frac{1}{2}u^{-1}\eth u\right)\right) = \int_{-E}^{E} k(s,s)ds, \tag{2.168}$$

where $k(s,t)$ is the kernel representing $\frac{1}{2}u^{-1}\eth u$. Its diagonal values are

$$k(s,s) = -\frac{1}{\pi}\frac{d\theta}{ds}, \tag{2.169}$$

where we use (2.139) and (2.146). Thus, the integral gives

$$\mathrm{Tr}\left(\mathbf{1}_{[-E,E]}\left(\frac{1}{2}u^{-1}\eth u\right)\right) = -\frac{2}{\pi}\theta(E) = -2(\langle N(E)\rangle - 1). \tag{2.170}$$

The remainder in the formula

$$\mathrm{Tr}(N_E\widehat{P}_\Lambda P_\Lambda) = 2E\frac{2\log\Lambda}{2\pi} - 2(\langle N(E)\rangle - 1) + r(E,\Lambda) \tag{2.171}$$

is therefore given by

$$r(E,\Lambda) = \frac{1}{2}\mathrm{Tr}(\mathbf{1}_{[-E,E]}u^{-1}\eth u\,\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}). \tag{2.172}$$

For $\Lambda\to\infty$ one has $\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}\to 0$ strongly. However, one needs to be a bit careful about the operator $T = \mathbf{1}_{[-E,E]}u^{-1}\eth u$, since it is unclear that it is of trace-class. One can check that the operator TT^* is a Hilbert-Schmidt operator. Also $\mathbf{1}_{[-E,E]}u^{-1}\eth u\, f$ is of trace-class, for any compactly supported function f, since θ is smooth. With f smooth, compactly supported and identically equal to 1 on $[-E,E]$ one has

$$\mathrm{Tr}(\mathbf{1}_{[-E,E]}u^{-1}\eth u\,\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}) = \mathrm{Tr}(\mathbf{1}_{[-E,E]}u^{-1}\eth u\,\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}\,f). \tag{2.173}$$

In fact, $\mathbf{1}_{[-E,E]}u^{-1}\eth u\,\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}$ is of trace-class by Lemma 2.21.
Moreover, the commutator $[\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}, f]$ is the conjugate of $\eth f$ by the function $s\mapsto\Lambda^{is}$. Thus, it is of trace-class and converges weakly to 0 as a family of Hilbert-Schmidt operators for $\Lambda\to\infty$. Since $\mathbf{1}_{[-E,E]}u^{-1}\eth u$ is Hilbert-Schmidt, we see that we can permute f with $\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}$ without affecting the limit. Since $\Pi_{[\frac{2\log\Lambda}{2\pi},\infty]}\to 0$ strongly, we obtain

$$r(E,\Lambda)\to 0 \quad \text{for} \quad \Lambda\to\infty. \tag{2.174}$$

This completes the proof of Theorem 2.18, making use of (2.66) to control the difference between Q_Λ and $\widehat{P}_\Lambda P_\Lambda$ in the statement.

REMARK 2.22. One can use Lemma 2.21 to estimate the angle of the projections Q_Λ and N_E. Indeed, we obtain

$$[N_E, \widehat{P}_\Lambda P_\Lambda] \sim [\mathbf{1}_{[-E,E]}, \frac{1}{2}u^{-1}\bar{d}u\,\Pi_{[-\infty,\frac{2\log\Lambda}{2\pi}]}] + [\mathbf{1}_{[-E,E]}, \Pi_{[0,\frac{2\log\Lambda}{2\pi}]}]. \tag{2.175}$$

The second commutator on the right hand side of (2.175) is of the order of $\log(E) + \log(\log(\Lambda))$ by the analysis of §3.2. The limit for $\Lambda \to \infty$ of the first commutator on the right hand side of (2.175) has Hilbert-Schmidt norm of the order of $\sqrt{\log(E)}$, as one gets from the estimate

$$\int_{|s|>E}\int_{t\in[-E,E]}\left|\frac{e^{2i\theta(s)} - e^{2i\theta(t)}}{s-t}\right|^2 dt\,ds = O(\log(E)). \tag{2.176}$$

6. The map $\mathfrak{E}$

Notice that the first term in (2.132) of Theorem 2.18 above, of the form $\frac{4E}{2\pi}\log\Lambda$, is in fact the symplectic volume $v(W(E,\Lambda))$ of the box

$$W(E,\Lambda) = \{(\lambda,s)\in\mathbb{R}_+^*\times\mathbb{R} \,:\, |\log\lambda|\le\log\Lambda\,,\ \text{and}\ |s|\le E\} \tag{2.177}$$

as in (2.46). The symplectic volume is computed in the symplectic space given by the product of the group $\mathbb{R}_+^* \sim \mathbb{R}$ by its dual $\mathbb{R}$ under the pairing $(\lambda, s)\mapsto\lambda^{is}$, with the symplectic form given by the product of the Haar measure by the dual one.
As we saw in the discussion of prolate functions in §3.3, the symplectic volume $v(W(E,\Lambda))$ gives a counting of the quantum states in $L^2(\mathbb{R}_+^*)$ subject to the infrared and ultraviolet cutoffs, respectively given by $|\log\lambda|\le\log\Lambda$ and $|s|\le E$. Theorem 2.18 leads one to expect the zeros of zeta to show up if one can relate our first "scaling" quantum system described in §3.2 to the regular representation of $\mathbb{R}_+^*$ on $L^2(\mathbb{R}_+^*)$ in such a way that the cutoffs match as

$$\begin{aligned} |q|\le\Lambda \quad &\mapsto \quad |\lambda|\le\Lambda \\ |p|\le\Lambda \quad &\mapsto \quad |\lambda|^{-1}\le\Lambda \\ |h(q,p)|\le E \quad &\mapsto \quad |s|\le E \end{aligned} \tag{2.178}$$

The map from even functions $f(u)$, for $u\in\mathbb{R}$, to their restriction $\lambda^{1/2}f(\lambda)$, for $\lambda\in\mathbb{R}_+^*$, is compatible with the first line of (2.178). In order to obtain the second line of (2.178), we should require that the sought for map $\mathfrak{E}$ fulfills the condition

$$\mathfrak{E}(f)(\lambda) = \mathfrak{E}(\hat{f})(\lambda^{-1})\,, \tag{2.179}$$

where $\hat{f} = \mathbf{F}_{e_\mathbb{R}}(f)$ is the Fourier transform relative to the basic character $e_\mathbb{R}$. This of course fails for the restriction map but it holds for a map $\mathfrak{E}$ of the form

$$\mathfrak{E}(f)(\lambda) = \lambda^{1/2}\sum_{n\in\mathbb{Z}} f(n\,\lambda)\,, \qquad \text{for}\quad \lambda\in\mathbb{R}_+^*. \tag{2.180}$$

This follows from the Poisson summation formula, provided we restrict to the codimension 2 subspace

$$\{f\in\mathcal{S}(\mathbb{R})\,:\ \ f(0) = \hat{f}(0) = 0\}\,. \tag{2.181}$$

REMARK 2.23. *The two boundary conditions of* (2.181) *exactly account for the additive* 2 *in Theorem 2.18.*

Since by construction we restrict to *even* functions f, we can ignore $n \le 0$ in the summation (up to a factor of 2) and rewrite (2.180) as

$$\mathfrak{E}(f)(\lambda) = \lambda^{1/2} \sum_{n=1}^{\infty} f(n\,\lambda)\,, \quad \text{for } \lambda \in \mathbb{R}_+^*. \tag{2.182}$$

The third line of (2.178) also matches, since the natural representation ϑ_{a} of $\mathbb{R}^*$ on $L^2(\mathbb{R})$, given by

$$(\vartheta_{\mathrm{a}}(\lambda) f)(x) = f(\lambda^{-1} x), \tag{2.183}$$

relates through $\mathfrak{E}$ to the regular representation ϑ_{m} of $\mathbb{R}^*$, by

$$\mathfrak{E}\,\vartheta_{\mathrm{a}}(\lambda) = |\lambda|^{1/2} \vartheta_{\mathrm{m}}(\lambda)\mathfrak{E}. \tag{2.184}$$

The map $\mathfrak{E}$ of (2.182) is not an isometry. Still, it gives a good indication of how to find the correct semi-classical (and quantum) system that refines Theorem 2.18 by the additional corrections associated to primes (cf. §7 below). Before we continue with the task of constructing the desired system, we can see that the zeros of the Riemann zeta function are already visible at this level, in the approximation of prolate functions by Hermite–Weber functions in the limit $\Lambda \to \infty$.

6.1. Hermite–Weber approximation and Riemann's ξ function.

We have mentioned in §3.3 the fact that, in the limit $\Lambda \to \infty$, the prolate spheroidal wave functions ψ_n recover the Hermite–Weber functions Ω_n (cf. [**271**]). The simplification of the prolate spheroidal wave functions into Hermite–Weber functions is used, in particular, in the theory of confocal resonators in laser optics.

Shortly after the first laser became operational in 1960, a wealth of results concurred in shaping the modern theory of optical resonators. In particular, the problem of describing explicitly the resonator eigenmodes was first approached numerically by Fox and Li [**133**], for both planar and confocal resonators. On the very same issue of the Bell System Technical Journal, another seminal paper on the subject by Boyd and Gordon [**31**] produced an exact analysis of the eigenmodes, in the case of a confocal resonator with square mirrors. This is one of the few cases where an exact solution can be given completely explicitly and it turns out that the eigenmodes are precisely the prolate spheroidal wave functions introduced by Pollak and Slepian in [**245**].

This follows very simply by computing the field on one of the two reflectors by the Huygens principle as a superposition of contributions from sources at the points of the other reflector surface. This results in an integral equation of the form

$$\psi(x) = \lambda \int_{-C}^{C} \psi(s)\, e^{ixs}\, ds,$$

which is the type of "finite Fourier transform" used in [**245**] to obtain the prolate spheroidal wave functions.

They observe that, if one is interested in the field in an area of the reflector which is mostly concentrated near the center, then one can efficiently replace the prolate spheroidal wave functions by Hermite–Weber functions, in agreement with the numerical results of [**133**].

In the problem we are now considering, we can follow a similar viewpoint. The exact shape of the cutoff region (2.41) is only relevant for its intersection with the region defined by the constraint $|h| \leq E$. Thus, if we replace the square

$$\{|q| \leq \Lambda \ \text{ and } \ |p| \leq \Lambda\}$$

by the disk $p^2+q^2 \leq \Lambda^2$, we do not alter the semi-classical calculations of §3.2 above. At the quantum level this amounts to replacing the projection Q_Λ by the projection onto the eigenspaces of the harmonic oscillator corresponding to the energy range $[0, \Lambda^2]$. The corresponding eigenfunctions are the Hermite–Weber functions Ω_n, with $2n + 1 \leq 2\pi\Lambda^2$. We restrict as above to even n to get even functions.
This disk cutoff is not suitable for our goal since it is not clear how to obtain the analog of the correspondence (2.178). However, we know that the prolate spheroidal functions ψ_n converge to the Ω_n for $\Lambda \to \infty$. The following proposition shows how the zeros of the Riemann zeta function are already present in this very rough approximation.
Recall from (2.7) that Riemann's ξ function (2.6) admits a description as

$$\xi(t) = \xi(0) \prod(1 - \frac{t^2}{\alpha^2}), \tag{2.185}$$

where $\rho = \frac{1}{2} + i\alpha$ runs through all the zeros of ζ with $\Im\,\rho > 0$.

PROPOSITION 2.24. *Let H_n^0 be the codimension two subspace of the linear span of the Ω_{2j}, for $j \leq n$, determined by the two linear conditions $f(0) = \hat{f}(0) = 0$. Let $\mathfrak{E}$ be the map defined in (2.182). Then the space of Fourier transforms of elements of $\mathfrak{E}(H_n^0)$ is the space of polynomial multiples $P(t)\,\xi(t)$, where $\xi(t)$ is Riemann's ξ function and P is a polynomial of degree $\leq n-2$.*

PROOF. By (2.6) one has,

$$\xi(t) = -\frac{1}{2}\left(\frac{1}{4} + t^2\right)\Gamma\left(\frac{1}{4} + \frac{it}{2}\right)\pi^{-\frac{1}{4}-\frac{it}{2}}\zeta\left(\frac{1}{2} + it\right). \tag{2.186}$$

With $\mathfrak{E}$ as in (2.182), we obtain

$$\mathfrak{E}(\Omega_0)(\lambda) = \lambda^{1/2}\vartheta(\lambda^2)\,, \tag{2.187}$$

where we have

$$\vartheta(x) := \sum_{n=1}^{\infty} e^{-n^2\pi x}. \tag{2.188}$$

The Fourier transform of $\mathfrak{E}(\Omega_0)$, which is of the form

$$\widehat{\mathfrak{E}(\Omega_0)}(s) = \int_0^\infty \mathfrak{E}(\Omega_0)(\lambda)\lambda^{is}\frac{d\lambda}{\lambda}\,, \tag{2.189}$$

is not well defined, since Ω_0 does not fulfill the boundary conditions of (2.181) and the integral diverges at 0. However, (2.189) makes sense for $\Re(is) > \frac{1}{2}$ and gives

$$\widehat{\mathfrak{E}(\Omega_0)}(s) = \frac{1}{2}\zeta^*\left(\frac{1}{2} + is\right)\,, \tag{2.190}$$

where ζ^* is given by (2.4), i.e.

$$\zeta^*(s) = \pi^{-s/2}\,\Gamma\left(\frac{s}{2}\right)\zeta(s). \tag{2.191}$$

Moreover, one has

$$\sum_{n=1}^{\infty} n^{2k} e^{-\pi n^2 y} = (-\pi)^{-k} (\frac{\partial}{\partial y})^k \vartheta(y)\,. \tag{2.192}$$

Thus one obtains for the function $f(x) = \pi^k x^{2k} e^{-\pi x^2}$ that

$$\mathfrak{E}(f)(\lambda) = \lambda^{1/2} \sum_1^{\infty} f(n\,\lambda) = \pi^k \lambda^{2k+\frac{1}{2}} \sum_{n=1}^{\infty} n^{2k} e^{-\pi n^2 \lambda^2}$$

hence, using (2.192) with $y = \lambda^2$,

$$\widehat{\mathfrak{E}(f)}(t) = \int_0^{\infty} \mathfrak{E}(f)(\lambda) \lambda^{it} \frac{d\lambda}{\lambda} = \frac{1}{2}(-1)^k \int_0^{\infty} y^{\frac{1}{4}+i\frac{t}{2}} y^k (\frac{\partial}{\partial y})^k \vartheta(y) \frac{dy}{y} \tag{2.193}$$

After integration by parts (2.193) gives

$$\widehat{\mathfrak{E}(f)}(t) = \frac{1}{2} \prod_{m=0}^{k-1} \left(\frac{1}{4} + i\frac{t}{2} + m\right) \zeta^* \left(\frac{1}{2} + it\right). \tag{2.194}$$

As above, this holds for $\Re(it) > \frac{1}{2}$. Using (2.179), we see that the Fourier transform of $\mathfrak{E}(f)$ is an entire function provided $f \in \mathcal{S}(\mathbb{R})$ fulfills the boundary conditions of (2.181). Thus, for

$$f(x) = \sum_{j=0}^{n} c_j\, \pi^j\, x^{2j}\, e^{-\pi x^2}, \qquad f \in H_n^0, \tag{2.195}$$

we obtain

$$\widehat{\mathfrak{E}(f)}(t) = -P(t)\,\xi(t), \tag{2.196}$$

where the polynomial $P(t)$ is the quotient

$$P(t) = \frac{\sum_{j=0}^{n} c_j \prod_{k=0}^{j-1} \left(\frac{1}{4} + i\frac{t}{2} + k\right)}{t^2 + \frac{1}{4}}. \tag{2.197}$$

Notice that the numerator of (2.197) is divisible by $t^2 + \frac{1}{4}$ exactly when f fulfills the boundary conditions of (2.181). $\square$

As explained above, this is an oversimplification in our case, since it does not fulfill the matching conditions (2.178). It is however useful, since the prolate spheroidal functions admit very efficient expansions in terms of Hermite-Weber functions, cf. e.g. [**129**].

If we were dealing with polynomials, Proposition 2.24 would yield the ideal J generated by ξ, and the zeros of ξ would be the spectrum of the quotient by J. This is exactly in line with seeing the zeros of zeta as an absorption spectrum, or as appearing in the cokernel of the map $\mathfrak{E}$ defined between suitable function spaces. An instance of such a realization using Sobolev spaces was given in [**71**] and will be explained in §9 Theorem 2.47. This type of result is mainly a motivation and admits many variants based on various choices of function spaces. For instance, for any $f \in \mathcal{S}(\mathbb{R})^{\text{even}}$ fulfilling the conditions of (2.181), we have the factorization

$$\widehat{\mathfrak{E}(f)}(t) = -Z(t)\,\xi(t)\,, \tag{2.198}$$

where $Z(t)$ is the entire function

$$Z(t) = \frac{\frac{1}{\Gamma(\frac{1}{4}+i\frac{t}{2})}\int_0^\infty u^{-\frac{3}{4}+i\frac{t}{2}}\, f(\sqrt{\frac{u}{\pi}})\, du}{t^2+\frac{1}{4}}. \tag{2.199}$$

7. The adèle class space: finitely many degrees of freedom

We now return to our goal, which is to describe a semi-classical system that accounts for all the additional terms indexed by prime numbers in (2.35). The space we construct is the configuration space (instead of the phase space) that accounts for the right terms in the formula (2.35). We first introduce this space, the *adèle class space*, as a quotient of the adèles of $\mathbb{Q}$ by the action of $\mathbb{Q}^*$. In Chapter 3 we give a simpler and more geometric description of the same space in terms of $\mathbb{Q}$-lattices modulo the relation of commensurability, and in Chapter 4 we give another, even more natural reinterpretation, in terms of Artin motives and their noncommutative generalization, the *endomotives* introduced in [**74**]. In this section we describe in detail the case of finitely many degrees of freedom, where one can give a complete proof of the trace formula of [**71**] (semi-local case).

DEFINITION 2.25. *Given topological spaces X_α and open subsets $U_\alpha \subset X_\alpha$, the restricted product of the X_α with respect to the U_α is the subspace of $\prod_\alpha X_\alpha$ of elements (x_α) such that $x_\alpha \in U_\alpha$ for all but finitely many α.*

One denotes this subspace by $\prod'_\alpha X_\alpha$ or by $\prod^{\rm res}_\alpha X_\alpha$, when the choice of the U_α is understood. It is by construction the union of the subsets

$$Y_S = \prod_{\alpha\in S} X_\alpha \times \prod_{\alpha\notin S} U_\alpha$$

and one endows $\prod'_\alpha X_\alpha$ with the topology for which a subset is open if and only if its intersection with all the above subsets Y_S is open. When each X_α is locally compact and almost all U_α are compact, the restricted product $\prod'_\alpha X_\alpha$ is locally compact.

The adèles of $\mathbb{Q}$ are the restricted product

$$\mathbb{A}_{\mathbb{Q}} = \prod_{v\in\Sigma_{\mathbb{Q}}}^{\rm res} \mathbb{Q}_v, \tag{2.200}$$

where the set $\Sigma_{\mathbb{Q}}$ is the set of valuations (places) of $\mathbb{Q}$, that is, $\Sigma_{\mathbb{Q}} = \mathcal{P}\cup\{\infty\}$, with $\mathcal{P}$ the set of primes in $\mathbb{Z}$. The field $\mathbb{Q}_v$ is the completion of $\mathbb{Q}$ in the valuation $|\cdot|_v$, that is, the p-adics $\mathbb{Q}_p$ at the finite places $p\in\mathcal{P}$ and $\mathbb{Q}_\infty = \mathbb{R}$ at the infinite (Archimedean) place $\infty\in\Sigma_{\mathbb{Q}}$. The restricted product is with respect to $\mathbb{Z}_p\subset\mathbb{Q}_p$.

One can equally describe the adèles of $\mathbb{Q}$ in the following way. First consider the profinite completion of $\mathbb{Z}$,

$$\hat{\mathbb{Z}} = \varprojlim_N \mathbb{Z}/N\mathbb{Z}. \tag{2.201}$$

The finite adèles of $\mathbb{Q}$ are given by

$$\mathbb{A}_{\mathbb{Q},f} = \hat{\mathbb{Z}}\otimes_{\mathbb{Z}}\mathbb{Q}, \tag{2.202}$$

and the full adèles (2.200) are then

$$\mathbb{A}_{\mathbb{Q}} = \mathbb{A}_{\mathbb{Q},f}\times\mathbb{R}. \tag{2.203}$$

The maximal compact subring of $\mathbb{A}_{\mathbb{Q},f}$ is identified with $\hat{\mathbb{Z}}$, which can be also be written in the form

$$\hat{\mathbb{Z}} = \prod_p \mathbb{Z}_p. \tag{2.204}$$

The idèles of $\mathbb{Q}$ are the invertible elements of $\mathbb{A}_{\mathbb{Q}}$, namely $\mathbb{A}_{\mathbb{Q}}^* = \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q}})$. (The topology on $\mathbb{A}_{\mathbb{Q}}^*$ is not the one induced by $\mathbb{A}_{\mathbb{Q}}$, but it is given by the embedding $x \mapsto (x, x^{-1})$.) Similarly, one considers $\mathbb{A}_{\mathbb{Q},f}^* = \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},f})$. The subgroup $\mathrm{GL}_1(\mathbb{Q})$ of $\mathbb{A}_Q^*$ is a discrete subgroup and the quotient group is the idèles class group

$$C_{\mathbb{Q}} = \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q}})/\mathrm{GL}_1(\mathbb{Q}). \tag{2.205}$$

One can then consider the group $C_{\mathbb{Q}}$ modulo the connected component of the identity $D_{\mathbb{Q}}$. The quotient is given by

$$C_{\mathbb{Q}}/D_{\mathbb{Q}} = \mathbb{A}_{\mathbb{Q},f}^*/\mathbb{Q}_+^* = \hat{\mathbb{Z}}^*, \tag{2.206}$$

the group of units (invertible elements of the ring) of $\hat{\mathbb{Z}}$. This identification will be useful when we consider in Chapter 3 the abelian class field theory of $\mathbb{Q}$. In fact, by the class field theory isomorphism, the group (2.206) is identified with the Galois group of the maximal abelian extension $\mathbb{Q}^{\mathrm{ab}}$ of $\mathbb{Q}$, namely with $\mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q}) = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})^{\mathrm{ab}}$, the abelianization of the absolute Galois group of $\mathbb{Q}$.

We can also consider the action of $\mathbb{Q}^*$ on the adèles $\mathbb{A}_{\mathbb{Q}}$. Unlike the action on idèles $\mathbb{A}_{\mathbb{Q}}^*$, the action on adèles is not "nice" in the usual topological sense. In other words, this is a typical case of a quotient that fails to be a nice topological space: the Hausdorff condition fails and typically there are no good functions on the quotient other than constants. The tools of noncommutative geometry are especially designed to treat this type of quotient in such a way that one can still use all the concepts of (smooth) geometry. The main point is that one replaces the ring of functions on the quotient (defined as functions invariant under the equivalence relation) with functions on the graph of the equivalence relation with a noncommutative convolution product dictated by the groupoid law of the equivalence relation. The *adèle class space* of $\mathbb{Q}$ is the quotient

$$X_{\mathbb{Q}} = \mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*, \tag{2.207}$$

regarded as a noncommutative space. In Chapter 4 we show how to use the tools of noncommutative geometry to give a cohomological interpretation of the spectral realization of the zeros of the Riemann zeta function , but in the present Chapter we shall give the direct description in more naive terms.

7.1. Geometry of the semi-local adèle class space.

In order to understand the adèle class space we first consider a semi-local version of this space, that is, one in which we restrict to finitely many degrees of freedom, by considering only finitely many places.

Let S be a finite set of places with $\infty \in S$. We consider the locally compact ring

$$\mathbb{A}_{\mathbb{Q},S} = \prod_{v \in S} \mathbb{Q}_v. \tag{2.208}$$

It contains $\mathbb{Q}$ as a subring using the diagonal embedding. We let $\mathbb{Q}_S$ denote the subring of $\mathbb{Q}$ given by rational numbers whose denominator only involves primes $p \in S$. In other words,

$$\mathbb{Q}_S = \{q \in \mathbb{Q} \,|\, |q|_v \leq 1\,,\ \forall v \notin S\}\,. \tag{2.209}$$

The group $\mathbb{Q}_S^*$ of invertible elements of the ring $\mathbb{Q}_S$ is of the form

$$\mathbb{Q}_S^* = \mathrm{GL}_1(\mathbb{Q}_S) = \{\pm p_1^{n_1} \cdots p_k^{n_k} \,:\, p_j \in S \setminus \{\infty\}\,,\, n_j \in \mathbb{Z}\}. \tag{2.210}$$

The semi-local adèle class space $X_{\mathbb{Q},S}$ is the quotient

$$X_{\mathbb{Q},S} := \mathbb{A}_{\mathbb{Q},S}/\mathbb{Q}_S^*. \tag{2.211}$$

The group

$$\mathbb{A}_{\mathbb{Q},S}^* = \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S}) = \prod_{p \in S} \mathrm{GL}_1(\mathbb{Q}_p) \tag{2.212}$$

acts naturally by multiplication on the quotient $X_{\mathbb{Q},S}$.

We shall show in §7.2 how to quantize this system, i.e. to define the Hilbert space $L^2(X_{\mathbb{Q},S})$ and to analyze the action of $\mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})$ on $L^2(X_{\mathbb{Q},S})$. For now we shall start with the simplest nontrivial examples where one can get a picture of the semi-local adèle class space and see in which sense it appears as a compactification of the following semi-local analogue of the idèle class group

$$C_{\mathbb{Q},S} = \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})/\mathbb{Q}_S^* \tag{2.213}$$

This group acts on $X_{\mathbb{Q},S}$ by multiplication and the orbit of $1 \in \mathbb{A}_{\mathbb{Q},S}$ (the unit of the ring) gives an embedding

$$C_{\mathbb{Q},S} \to X_{\mathbb{Q},S}. \tag{2.214}$$

Notice first that, if the set S consists solely of the point ∞, then the system we obtain is the one we have already discussed in Sections 3 and 5.1 above. We now describe concretely the space $X_{\mathbb{Q},S}$ in other simple cases where S consists of two or three places.

7.1.1. *Two places:* $S = \{p, \infty\}$.

We let p be a prime number and take $S = \{p, \infty\}$. We get

$$\begin{aligned} \mathbb{A}_{\mathbb{Q},S} &= \mathbb{Q}_p \times \mathbb{R} \\ \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S}) &= \mathbb{Q}_p^* \times \mathbb{R}^* \\ \mathbb{Q}_S^* &= \{\pm p^n\,,\, n \in \mathbb{Z}\} \end{aligned}$$

Thus in this case the space $X_{\mathbb{Q},S}$ is the quotient of $\mathbb{Q}_p \times \mathbb{R}$ by the group $\mathbb{Q}_S^*$ generated by the symmetry $\sigma(x, y) = (-x, -y)$ and the scaling transformation s_p

$$s_p(x, y) = \,(p\,x,\ p\,y)\,.$$

Both of these transformations preserve the function

$$\tilde{f}(x, y) = \,|x|_p\,|y|_\infty \in \mathbb{R}_+ \tag{2.215}$$

which can be viewed as a continuous function

$$f \,:\, X_{\mathbb{Q},S} \to \mathbb{R}_+\,. \tag{2.216}$$

The following proposition shows in which sense $X_{\mathbb{Q},S}$ can be thought of as a compactification of $C_{\mathbb{Q},S}$.

PROPOSITION 2.26. *The subspace $C_{\mathbb{Q},S} \subset X_{\mathbb{Q},S}$ is the inverse image of $\mathbb{R}_+^*$ by f. The fiber $f^{-1}(0)$ is the union of three pieces,*

$$f^{-1}(0) = T_p \cup T_\infty \cup \{0\},$$

where $T_p \sim \mathbb{R}_+^*/p^{\mathbb{Z}}$ and $T_\infty \sim \mathbb{Z}_p^*/\{\pm 1\}$ while $\{0\}$ is a single point.

PROOF. The first statement follows from

$$\tilde{f}(x,y) \neq 0 \Leftrightarrow (x,y) \in \mathbb{Q}_p^* \times \mathbb{R}^* \subset \mathbb{Q}_p \times \mathbb{R}.$$

Note that the fiber of f over any nonzero $\varepsilon \in \mathbb{R}_+$ is easily identified with $\mathbb{Z}_p^*$,

$$f^{-1}(\varepsilon) \sim \mathbb{Z}_p^*, \quad \forall \varepsilon \neq 0$$

where one can for instance use the fundamental domain

$$D = \mathbb{Z}_p^* \times \mathbb{R}_+^* \tag{2.217}$$

for the action of $\mathbb{Q}_S^*$ on $\mathbb{Q}_p^* \times \mathbb{R}^*$ to obtain the required isomorphism.
The fiber $f^{-1}(0)$ of f over the point $\varepsilon = 0$ is no longer $\mathbb{Z}_p^*$ and is the union of three pieces

$$f^{-1}(0) = T_p \cup T_\infty \cup \{0\}$$

corresponding respectively to

- $x = 0$, $y \neq 0$ which gives a torus $T_p \sim \mathbb{R}_+^*/p^{\mathbb{Z}}$ using the map $(x,y) \to |y|$.
- $x \neq 0$, $y = 0$ which gives the compact $T_\infty \sim \mathbb{Z}_p^*/\{\pm 1\}$ as quotient of $\mathbb{Q}_p^*$ by the action of $\mathbb{Q}_S^*$.
- $x = 0$, $y = 0$ which gives one point 0.

$\square$

Let us briefly describe the quotient topology on $X_{\mathbb{Q},S}$. On the open set which is the complement of $f^{-1}(0)$ it coincides with the locally compact topology of $C_{\mathbb{Q},S}$. The convergence of the fiber $f^{-1}(\varepsilon) \sim \mathbb{Z}_p^*$ to $f^{-1}(0)$ is governed by

PROPOSITION 2.27. *Let $P_n \in C_{\mathbb{Q},S}$ and (x_n, y_n) be the coordinates of the unique representative of P_n in the fundamental domain D (of (2.217)). Then, in the (non-Hausdorff) quotient topology,*

- *$P_n \to 0 \Leftrightarrow f(P_n) \to 0$.*
- *For $x \in T_\infty \sim \mathbb{Z}_p^*/\{\pm 1\}$ one has $P_n \to x \Leftrightarrow f(P_n) \to 0$ and $x_n \to x$ in $\mathbb{Z}_p^*/\{\pm 1\}$.*
- *For $y \in T_p \sim \mathbb{R}_+^*/p^{\mathbb{Z}}$ one has $P_n \to y \Leftrightarrow f(P_n) \to 0$ and $y_n \to y$ in $\mathbb{R}_+^*/p^{\mathbb{Z}}$.*

PROOF. Since f is continuous one has $P_n \to 0 \Rightarrow f(P_n) \to 0$. Conversely, since, for $P_n = (x_n, y_n) \in D$, one has $f(P_n) = |y_n|$ one gets $|y_n| \to 0$ and one can find a sequence p^{k_n} of positive powers of p such that $k_n \to \infty$ and $p^{k_n}|y_n| \to 0$. Then the two components of $s_p^{k_n}(x_n, y_n) = (p^{k_n}x_n, p^{k_n}y_n)$ converge to 0 which shows that $P_n \to 0$ in the quotient topology.

Let $x \in T_\infty \sim \mathbb{Z}_p^*/\{\pm 1\}$. One has $f(x) = 0$; thus $P_n \to x \Rightarrow f(P_n) \to 0$. In fact if $P_n \to x$ there exists a sequence p^{k_n} of powers of p such that $(p^{k_n}x_n, p^{k_n}y_n) \to (x, 0)$ in $\mathbb{Q}_p \times \mathbb{R}$. But this implies that $k_n = 0$ for n large enough and thus that $x_n \to x$ in

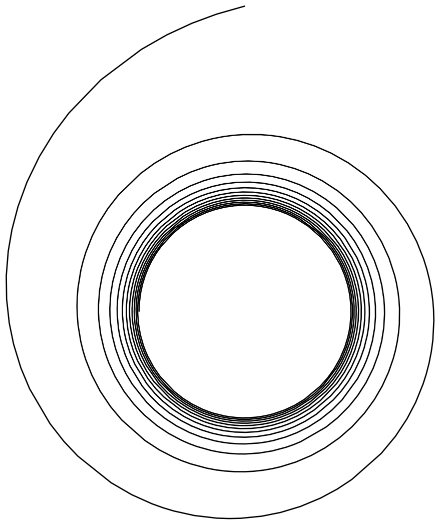


FIGURE 9. Holonomy

$\mathbb{Z}_p^*/\pm 1$. Conversely if $x_n \to x$ in $\mathbb{Z}_p^*/\{\pm 1\}$ and $f(P_n) \to 0$ then $(x_n, y_n) \to (x, 0)$ in $\mathbb{Q}_p \times \mathbb{R}$ so that $P_n \to x$.

Let now $y \in T_p \sim \mathbb{R}_+^*/p^{\mathbb{Z}}$. One has $f(y) = 0$; thus $P_n \to y \Rightarrow f(P_n) \to 0$. If $P_n \to y$ there exists a sequence p^{k_n} of powers of p such that $(p^{k_n} x_n, p^{k_n} y_n) \to (0, y)$ in $\mathbb{Q}_p \times \mathbb{R}$. This implies that $k_n \to \infty$ and $p^{k_n} y_n \to y$ so that in the quotient circle $\mathbb{R}_+^*/p^{\mathbb{Z}}$ one has $y_n \to y$. Conversely if $y_n \to y$ in $\mathbb{R}_+^*/p^{\mathbb{Z}}$ and $f(P_n) \to 0$, one can find a sequence p^{k_n} of positive powers of p such that $k_n \to \infty$ and $p^{k_n} y_n \to y$. It follows that $(p^{k_n} x_n, p^{k_n} y_n) \to (0, y)$ in $\mathbb{Q}_p \times \mathbb{R}$. $\square$

Proposition 2.27 shows that the quotient topology on the space $X_{\mathbb{Q},S}$ is non-Hausdorff and, when $y_n \to 0$, the same sequence $(x_n, y_n) \in D$ has three different limit points in the boundary fiber $f^{-1}(0)$,

- $\lim x_n \in T_\infty$
- $\lim y_n \in T_p$
- $0 \in X_{\mathbb{Q},S}$

The convergence to $T_p \sim \mathbb{R}_+^*/p^{\mathbb{Z}}$ is very suggestive of holonomy in the framework of foliations and, using a logarithmic scale,

$$\mathbb{R}_+^*/p^{\mathbb{Z}} \sim \mathbb{R}/(\mathbb{Z} \log p)$$

it corresponds to what happens in the limit cycle of the foliation associated to a flow as in Figure 9. Indeed the conditions $y_n \to 0$ *and* $y_n \to y$ in $\mathbb{R}_+^*/p^{\mathbb{Z}}$ exactly mean that the points $-\log y_n$ converge to the point $-\log y$ in the limit cycle $\mathbb{R}/(\mathbb{Z} \log p)$.

In fact one can be more precise and see what happens by analyzing the C^*-algebras involved. The C^*-algebra which encodes the topology of the quotient space $X_{\mathbb{Q},S}$ is the crossed product

$$\mathcal{A} = C_0(\mathbb{A}_{\mathbb{Q},S}) \rtimes \mathbb{Q}_S^*\,, \tag{2.218}$$

and the C^*-algebra associated to the boundary fiber $f^{-1}(0)$ is by construction the crossed product

$$A = C_0(\tilde{f}^{-1}(0)) \rtimes \mathbb{Q}_S^*\,. \tag{2.219}$$

The decomposition of $\tilde{f}^{-1}(0)$ as the union

$$\tilde{f}^{-1}(0) = V_\infty \cup V_p\,, \quad V_\infty = \mathbb{Q}_p \times 0\,, V_p = 0 \times \mathbb{R}$$

of the V_v, $v \in \{p, \infty\}$, over their common point 0 is $\mathbb{Q}_S^*$-equivariant. Thus one gets that the C^*-algebra A is the fiber product of the C^*-algebras A_v,

$$A_v = C_0(V_v) \rtimes \mathbb{Q}_S^*\,, \quad v \in \{p, \infty\} \tag{2.220}$$

over the augmentation morphisms $\epsilon_v : A_v \to C^*(\mathbb{Q}_S^*)$ given by evaluation at 0. Moreover for each of the A_v one has an exact sequence of the form

$$0 \to C(T_v) \otimes \mathcal{K} \to A_v \stackrel{\epsilon_v}{\to} C^*(\mathbb{Q}_S^*) \to 0, \tag{2.221}$$

where $\mathcal{K}$ is the algebra of compact operators. This follows using the fact that the action of $\mathbb{Q}_S^*$ on $\mathbb{R}^*$ or $\mathbb{Q}_p^*$ is free. It follows in particular that, in the case of two places, the C^*-algebra $\mathcal{A}$ of (2.218) is of type I (cf. [**120**]), i.e. that all factor representations of $\mathcal{A}$ generate a type I factor. This no longer holds in general, as we shall see now with a specific example.

7.1.2. *Three places $S = \{p, q, \infty\}$.*

We give an example to show that, as soon as the system is constructed out of a set S that contains at least two finite primes in addition to the point ∞, then the quotient $X_{\mathbb{Q},S}$ should be regarded as a noncommutative space.

LEMMA 2.28. *Consider the set of 3 places $S = \{p, q, \infty\}$. Then*

- *The quotient $X_{\mathbb{Q},S}$ is not countably separated.*
- *The crossed product C^*-algebra*

$$\mathcal{A} = C_0(\mathbb{A}_{\mathbb{Q},S}) \rtimes \mathbb{Q}_S^*$$

is not of type I.

PROOF. The irrationality of $\log p/\log q$ when p and q are distinct primes shows that the subgroup

$$P = \{\pm p^n q^m \,:\, n, m \in \mathbb{Z}\} \tag{2.222}$$

is dense in $\mathbb{R}^*$. It follows that $\mathbb{Q}_S^*$ acts ergodically on $\{0\} \times \mathbb{R} \subset \mathbb{A}_{\mathbb{Q},S}$ and thus the quotient $X_{\mathbb{Q},S}$ is not countably separated.
Since the crossed product C^*-algebra admits $C_0(\mathbb{R}) \rtimes P$ as a quotient, it has non type I factor representations and hence fails to be of type I. □

Taking $p = 2$, $q = 3$ as the simplest example it is amusing to note that it is the irrationality of the ratio $\log 3/\log 2$ that is responsible for the noncommutative nature of the quotient $\mathbb{A}_{\mathbb{Q},S}/\mathbb{Q}_S^*$ corresponding to the three places $S = \{2, 3, \infty\}$. The irrationality of $\log 3/\log 2$ has a manifestation in the theory of western classical music, where a frequency ratio of 2 corresponds to passing to the octave and the perfect fifth is the frequency ratio 3/2. The fact that the ratio $\log 3/\log 2$ is only approximated by the rational number 19/12 is responsible for the difference between the "circulating temperament" of Baroque music (e.g. the Well Tempered Clavier) and the "equal temperament" of XIX century music.

7.2. The Hilbert space $L^2(X_S)$ and the trace formula.

Our purpose in this section is to quantize the system given by the action of $\mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})$ on the semi-local adèle class space $X_{\mathbb{Q},S}$, i.e. to define the Hilbert space $L^2(X_{\mathbb{Q},S})$ and to analyze the action of $\mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})$ on $L^2(X_{\mathbb{Q},S})$. The additive group of the ring $\mathbb{A}_{\mathbb{Q},S}$ is a locally compact abelian group and we can consider the corresponding Haar measure, which we denote by da. The module $|\lambda|_S$ is defined by $d(\lambda a) = |\lambda|\, da$ and is given by the product

$$|\lambda|_S = \prod_{v\in S} |\lambda_v|, \quad \forall \lambda = (\lambda_v) \in \prod_{v\in S} \mathbb{Q}_v \tag{2.223}$$

of the modules $|\lambda_v|$ of the components of λ.
For any $\lambda \in \mathbb{Q}_S^*$ one has $|\lambda|_S = 1$, since $\mathbb{Q}_S^*$ is the subgroup of $\mathrm{GL}_1(\mathbb{Q})$ defined by

$$\mathbb{Q}_S^* = \{\gamma \in \mathrm{GL}_1(\mathbb{Q}) \,:\, |\gamma|_v = 1\,,\ \forall v \notin S\}. \tag{2.224}$$

We now construct the Hilbert space associated to the quotient $X_{\mathbb{Q},S}$. We need the following preliminary results.

Lemma 2.29. *Let $f \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$. The series*

$$\mathbf{\Sigma}(f)(x) = \sum_{q\in\mathbb{Q}_S^*} f(qx) \tag{2.225}$$

is absolutely convergent for any $x \in \mathbb{A}_{\mathbb{Q},S}^*$.
One has, with $m = \mathrm{Card}\, S$,

$$|\mathbf{\Sigma}(f)(x)| = O(|\log|x||^{m-1})\,, \quad \textit{for } |x| \to 0\,, \tag{2.226}$$

and $|\mathbf{\Sigma}(f)(x)|$ is of rapid decay in $|x|$ for $|x| \to \infty$.

Proof. Let us consider the map from $\mathbb{A}_{\mathbb{Q},S}^*$ to $\mathbb{R}^S$ given by

$$u \in \prod_S \mathbb{Q}_v^* \mapsto \ell(u) = (\log|u_v|)_v \in \mathbb{R}^S\,. \tag{2.227}$$

The group $\mathbb{Q}_S^*$ is a discrete cocompact subgroup of the kernel of the map $\lambda \mapsto |\lambda|_S$ from $\mathbb{A}_{\mathbb{Q},S}^*$ to $\mathbb{R}_+^*$. The kernel of the restriction of ℓ to $\mathbb{Q}_S^*$ is $\{q \in \mathbb{Q}^* \,|\, |q|_v = 1, \forall v \in \Sigma_{\mathbb{Q}}\}$ which is ± 1. The image $L = \ell(\mathbb{Q}_S^*)$ by the map ℓ is a lattice in the hyperplane

$$H = \{y = (y_v) \in \mathbb{R}^S \,|\, \sum y_v = 0\}\,. \tag{2.228}$$

For any element of H one has, with $m = \mathrm{Card}\, S$,

$$\sum |y_v| \le 2m \sup y_v\,. \tag{2.229}$$

Indeed one has $\sum_{y_v>0} |y_v| = \sum_{y_v<0} |y_v|$ and $\frac{1}{2}\sum |y_v| = \sum_{y_v>0} y_v \le m \sup y_v$.
Let $n \in \mathbb{N}$, and

$$B_n = \{x \in \mathbb{A}_{\mathbb{Q},S} \,|\, |x_v| \le n\,, \quad \forall v \in S\}\,. \tag{2.230}$$

Given $f \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$ there exists a sequence $\lambda_n > 0$ of rapid decay such that

$$|f(x)| \le \sum \lambda_n\, \mathbf{1}_{B_n}(x)\,. \tag{2.231}$$

Using (2.231) the first statement of the lemma follows from the following estimate. There exist constants $0 < c_j < \infty$ such that, for any $x \in \mathbb{A}^*_{\mathbb{Q},S}$,

$$|\mathbf{\Sigma}(\mathbf{1}_{B_n})(x)| \leq c_1 |c_2 + m \log n - \log |x||^{m-1} . \tag{2.232}$$

Let us prove (2.232). By construction $|\mathbf{\Sigma}(\mathbf{1}_{B_n})(x)|$ is the number of elements $q \in \mathbb{Q}^*_S$ such that $|q_v x_v| \leq n$ for all $v \in S$. Up to a finite factor (equal to 2) it is, using the map ℓ of (2.227), the cardinality of

$$Y = \{y = (y_v) \in L \mid y_v + \log |x_v| \leq \log n, \forall v \in S\} . \tag{2.233}$$

Since L is a lattice in H there exists $\epsilon > 0$ such that the cubes of size ϵ centered at points of L are pairwise disjoint. Thus the cardinality of Y is less than ϵ^{-m+1} times the volume in H of the subset

$$Y' = \{y = (y_v) \in H \mid y_v + \log |x_v| \leq \log n + \epsilon, \forall v \in S\} . \tag{2.234}$$

The defining inequalities for this subset of H are of the form $y_v \leq \alpha_v$ and, in the variables $z_v = \alpha_v - y_v$, one is dealing with the simplex (with $\alpha = \sum \alpha_v$)

$$S = \{z = (z_v) \mid \sum z_v = \alpha\,, \quad z_v \geq 0, \forall v \in S\} , \tag{2.235}$$

whose volume is proportional to α^{m-1}. Here $\alpha_v = \log n + \epsilon - \log |x_v|$ and $\alpha = \sum \alpha_v = m(\log n + \epsilon) - \log |x|$; thus one gets (2.232).
To get the estimate (2.226) for $|x| \to 0$ one can use the convexity of the map $x \mapsto x^a$ for $a = m - 1$, i.e. the inequality $(\frac{u+v}{2})^a \leq \frac{u^a+v^a}{2}$ (for $u, v > 0$), to replace (2.232) for $|x| < 1$ by

$$|\mathbf{\Sigma}(\mathbf{1}_{B_n})(x)| \leq c'_1 (c_2 + m \log n)^{m-1} + c'_1 |\log |x||^{m-1} . \tag{2.236}$$

Since the sequence λ_n is of rapid decay one then gets

$$|\mathbf{\Sigma}(f)(x)| \leq c_3 + c_4 |\log |x||^{m-1} \tag{2.237}$$

where $c_3 = \sum_n c'_1 (c_2 + m \log n)^{m-1} \lambda_n < \infty$ and $c_4 = c'_1 \sum \lambda_n < \infty$.
Next we consider the behaviour of $|\mathbf{\Sigma}(f)(x)|$ for $|x| \to \infty$. For $|x| > 1$ one has from (2.232) that $|\mathbf{\Sigma}(\mathbf{1}_{B_n})(x)| \leq c_1 (c_2 + m \log n)^{m-1}$ and moreover $|\mathbf{\Sigma}(\mathbf{1}_{B_n})(x)| = 0$ for $|x| > n^m$ since one has $|y| \leq n^m$ for all $y \in B_n$. Thus for $|x| > n_0^m$ one gets

$$|\mathbf{\Sigma}(f)(x)| \leq \sum_{n_0}^{\infty} \lambda_n \, c_1 (c_2 + m \log n)^{m-1}$$

which is of rapid decay in n_0, so that $|\mathbf{\Sigma}(f)(x)|$ is of rapid decay in $|x|$. □

The quotient group

$$C_{\mathbb{Q},S} := \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})/\mathbb{Q}^*_S \tag{2.238}$$

is a locally compact modulated group (Definition 2.8) with module

$$\mathrm{Mod}(\lambda) := |\lambda|_S, \quad \forall \lambda \in C_{\mathbb{Q},S} \tag{2.239}$$

given by (2.223). The kernel $C^1_{\mathbb{Q},S}$ of the module is a compact subgroup of $C_{\mathbb{Q},S}$ and $C_{\mathbb{Q},S}$ is (non-canonically) isomorphic to $C^1_{\mathbb{Q},S} \times \mathbb{R}^*_+$.
We normalize the multiplicative Haar measure $d^*\lambda$ of $C_{\mathbb{Q},S}$ by

$$\int_{|\lambda| \in [1,\Lambda]} d^*\lambda \sim \log \Lambda \quad \text{when} \quad \Lambda \to \infty. \tag{2.240}$$

PROPOSITION 2.30. *The following defines a map $\mathfrak{E}$ with dense range from $\mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$ to $L^2(C_{\mathbb{Q},S})$:*

$$\mathfrak{E}(f)(x) = |x|^{1/2}\mathbf{\Sigma}(f)(x)\,, \quad \forall x \in C_{\mathbb{Q},S}\,. \tag{2.241}$$

PROOF. By construction $\mathbf{\Sigma}(f)(x)$ only depends upon the class of x in $C_{\mathbb{Q},S}$. Moreover by Lemma 2.29 the function $\mathfrak{E}(f)(x)$ is of exponential decay on $C_{\mathbb{Q},S}$ in terms of $|\log|x||$. Thus it is square integrable. Let D be a fundamental domain for the action of $\mathbb{Q}_S^*$ in $\prod \mathbb{Q}_v^*$. One can for instance take $D = \prod_{p\in S}\mathbb{Z}_p^* \times \mathbb{R}_+^*$. Then for f smooth with compact support in D one has $\mathfrak{E}(f)(x) = |x|^{1/2}f(x)$ and such functions are dense in $L^2(C_{\mathbb{Q},S})$. □

Thus, it makes sense to give the following definition.

DEFINITION 2.31. *We define the Hilbert space $L^2(X_{\mathbb{Q},S})$ as the completion of the Bruhat-Schwartz space $\mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$ with respect to the inner product*

$$\langle f_1, f_2\rangle_S = \langle \mathfrak{E}(f_1), \mathfrak{E}(f_2)\rangle\,. \tag{2.242}$$

We now give an equivalent formulation of the inner product. We let $L^2(\mathbb{A}_{\mathbb{Q},S})$ be the Hilbert space of square integrable functions on $\mathbb{A}_{\mathbb{Q},S}$. It is the completion of the Bruhat-Schwartz space $\mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$ with respect to the standard inner product

$$\langle f_1,\, f_2\rangle := \int_{\mathbb{A}_{\mathbb{Q},S}} \overline{f_1(a)}\, f_2(a)\, da. \tag{2.243}$$

The scaling operator $\vartheta_{\mathrm{a}}(\lambda)$, for $\lambda \in \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})$, is defined as

$$(\vartheta_{\mathrm{a}}(\lambda)\,\xi)(a) = \xi(\lambda^{-1}a), \quad \forall a \in \mathbb{A}_{\mathbb{Q},S},\ \forall \xi \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S}). \tag{2.244}$$

This fullfills the property that

$$\vartheta_{\mathrm{a}}(\lambda)^*\vartheta_{\mathrm{a}}(\lambda) = \vartheta_{\mathrm{a}}(\lambda)\vartheta_{\mathrm{a}}(\lambda)^* = |\lambda|_S. \tag{2.245}$$

LEMMA 2.32. *For any choice of $f_i \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$, $i = 1, 2$, the series*

$$\sum_{\gamma\in\mathbb{Q}_S^*} \langle f_1, \vartheta_{\mathrm{a}}(\gamma)f_2\rangle \tag{2.246}$$

of inner products in $L^2(\mathbb{A}_{\mathbb{Q},S})$ is absolutely convergent. The sum of the series coincides with the positive definite inner product $\langle f_1, f_2\rangle_S$ on $\mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$.

PROOF. Let $\mathbf{1}_{B_n}$ be the characteristic function of the set B_n of (2.230). Then, for any rational number of the form

$$\gamma = \pm\frac{x_1}{x_2} \quad \text{with} \quad x_j \in \mathbb{N}, \tag{2.247}$$

such that $\gamma \in \mathbb{Q}_S^*$, one has (up to a fixed normalization constant)

$$\langle \mathbf{1}_{B_n}, \vartheta_{\mathrm{a}}(\gamma)\,\mathbf{1}_{B_n}\rangle = n^m(\max\{x_j\})^{-1}. \tag{2.248}$$

Thus, using (2.231), the series (2.246) converges like the geometric series $e^{-h(\gamma)}$, where the height $h(\gamma) = \log\max\{x_j\}$ is equivalent to a length function on the finitely generated abelian group $\mathbb{Q}_S^*$. This is sufficient to get the required convergence.

Before showing the second part, it is worthwhile to check the positivity directly. Notice that, for any $f \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$, the function

$$\gamma \mapsto \phi(\gamma) = \langle f, \vartheta_{\mathrm{a}}(\gamma)f\rangle \tag{2.249}$$

is the coefficient of a unitary representation and hence is of positive type. By the above it is in $\ell^1(\mathbb{Q}_S^*)$, so that its sum $\sum_\gamma \phi(\gamma)$ is positive ([**120**], Proposition 18.3.6). Finally to obtain the equality (valid modulo a normalization constant)

$$\sum_{\gamma \in \mathbb{Q}_S^*} \langle f_1, \vartheta_{\mathrm{a}}(\gamma) f_2 \rangle = \langle f_1, f_2 \rangle_S\,, \quad \forall f_j \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S})\,, \tag{2.250}$$

one applies Fubini's Theorem to the double sum

$$\overline{\mathfrak{E}(f_1)(x)}\mathfrak{E}(f_2)(x) = |x| \sum_r \sum_s \bar{f}_1(rx) f_2(sx) = |x| \sum_\gamma \sum_r \bar{f}_1(rx)\vartheta_{\mathrm{a}}(\gamma) f_2(rx)$$

together with the equality, valid up to a normalization constant,

$$\int_{C_{\mathbb{Q},S}} \mathbf{\Sigma}(f)(\lambda)|\lambda| d^*\lambda = \int_{\mathbb{A}_{\mathbb{Q},S}} f(a) da\,. \tag{2.251}$$

$\square$

The radical of the inner product $\langle f_1, f_2 \rangle_S$ of (2.246) contains the linear span of functions of the form

$$f - \vartheta_{\mathrm{a}}(\gamma)\, f, \quad \forall f \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S}), \forall \gamma \in \mathbb{Q}_S^*. \tag{2.252}$$

This shows that it automatically reduces $\mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$ to its co-invariants under the action of $\mathbb{Q}_S^*$.
This fact is exactly analogous to what happens in the context of constructive gauge theories on lattices, where two functions of field configurations that are related by a gauge transformation differ only by an element of the radical of the inner product defining the physical Hilbert space (see [**261**] Theorem 2.1 and p.17).

The equality (2.244) defines the action of $\mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})$ on the Hilbert space $L^2(X_{\mathbb{Q},S})$ and (2.245) still holds. Moreover, this action is the identity on $\mathbb{Q}_S^* = \mathrm{GL}_1(\mathbb{Q}_S) \subset \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},S})$ and given a smooth compactly supported function h on $C_{\mathbb{Q},S}$, it makes sense to define

$$\vartheta_{\mathrm{a}}(h) = \int_{C_{\mathbb{Q},S}} h(g)\vartheta_{\mathrm{a}}(g)\, dg \tag{2.253}$$

as an operator acting on $L^2(X_S)$.

We now show that the Fourier transform $\mathbf{F}$ on $\mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$ extends to a unitary operator on the Hilbert space $L^2(X_S)$. For each place $v \in S$ we choose a basic character α_v of $\mathbb{Q}_v$. This makes it possible to identify the locally compact abelian group $\mathbb{Q}_v$ with its Pontrjagin dual by the pairing

$$\langle a, b\rangle := \alpha_v(ab), \quad \forall a, b \in \mathbb{Q}_v. \tag{2.254}$$

We normalize the additive Haar measure so that it is self-dual. Then $\alpha = \prod_v \alpha_v$ is a basic character of the additive group $\mathbb{A}_S$, and we let $\mathbf{F}_\alpha$ denote the corresponding Fourier transform, acting on the Bruhat-Schwartz space $\mathcal{S}(\mathbb{A}_S)$.

Lemma 2.33. *Consider the character $\alpha = \prod_v \alpha_v$ and the corresponding Fourier transform $\mathbf{F}_\alpha$ as above. The map $f \mapsto \mathbf{F}_\alpha(f)$, for $f \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S})$, extends uniquely to a unitary operator $\underline{\mathbf{F}}_\alpha$ on the Hilbert space $L^2(X_S)$.*

PROOF. One uses (2.96) and $|\gamma| = 1$ for $\gamma \in \mathbb{Q}_S^*$ to get the equalities

$$\langle \mathbf{F}_\alpha f, \mathbf{F}_\alpha f\rangle = \langle f, f\rangle \quad \text{and} \quad \mathbf{F}_\alpha(\vartheta_{\mathrm{a}}(\gamma)f) = \vartheta_{\mathrm{a}}(\gamma^{-1})\mathbf{F}_\alpha(f). \tag{2.255}$$

One then has, by (2.250),

$$\begin{aligned}\langle \mathbf{F}_\alpha f_1, \mathbf{F}_\alpha f_2\rangle_S &= \sum_{\gamma\in\mathbb{Q}_S^*} \langle \mathbf{F}_\alpha f_1, \vartheta_{\mathrm{a}}(\gamma)\mathbf{F}_\alpha f_2\rangle = \sum_{\gamma\in\mathbb{Q}_S^*} \langle \mathbf{F}_\alpha f_1, \mathbf{F}_\alpha \vartheta_{\mathrm{a}}(\gamma^{-1}) f_2\rangle \\ &= \sum_{\gamma\in\mathbb{Q}_S^*} \langle f_1, \vartheta_{\mathrm{a}}(\gamma^{-1}) f_2\rangle = \langle f_1, f_2\rangle_S \,.\end{aligned}$$

□

We let w be the unitary identification of $L^2(X_{\mathbb{Q},S})$ with $L^2(C_{\mathbb{Q},S})$ provided by Proposition 2.30. It intertwines the representation ϑ_{a} of (2.253) with the regular representation ϑ_{m} of $C_{\mathbb{Q},S}$

$$(\vartheta_{\mathrm{m}}(\lambda)\,\xi)(v) := \xi(\lambda^{-1}v)\,, \quad \forall \xi \in L^2(C_{\mathbb{Q},S})\,, \tag{2.256}$$

by the equality

$$w\,\vartheta_{\mathrm{a}}(\lambda) w^{-1} = |\lambda|^{1/2}\,\vartheta_{\mathrm{m}}(\lambda)\,, \quad \forall \lambda \in C_{\mathbb{Q},S}\,. \tag{2.257}$$

We now prove the analogue of Lemma 2.20. For each place v, the multiplicative group $\mathbb{Q}_v^*$ is a modulated group, and by [**281**] there exists a unitary function $u_v \in L^\infty(\widehat{\mathbb{Q}_v^*})$ such that, for any character χ of $\mathbb{Q}_v^*$

$$\int_{\mathbb{Q}_v^*} \mathbf{F}_{\alpha_v}(f)(x)\chi(x^{-1})|x|^{1/2} d^*x = u_v(\chi) \int_{\mathbb{Q}_v^*} f(x)\chi(x)|x|^{1/2} d^*x \tag{2.258}$$

where $\mathbf{F}_{\alpha_v}$ is the Fourier transform relative to the character α_v. By [**281**], the function u_v is given by the ratio of the local factors of L-functions.
The dual of the embedding $\mathbb{Q}_v^* \hookrightarrow C_S$ given by $u \mapsto (1, 1, \ldots, u, \ldots, 1)$ is a projection $\pi_v : \widehat{C_{\mathbb{Q},S}} \to \widehat{\mathbb{Q}_v^*}$.

LEMMA 2.34. *On $L^2(X_{\mathbb{Q},S})$ one has*

$$\underline{\mathbf{F}}_\alpha = w^{-1} \circ I \circ \mathbf{F}_C^{-1} \circ u \circ \mathbf{F}_C \circ w, \tag{2.259}$$

where u is the multiplication operator by the function

$$u := \prod_S u_v \circ \pi_v \,. \tag{2.260}$$

PROOF. The composition $I \circ w \circ \underline{\mathbf{F}}_\alpha \circ w^{-1}$ is a unitary operator in $L^2(C_{\mathbb{Q},S})$ which commutes with the regular representation ϑ_{m} of $C_{\mathbb{Q},S}$ (2.256). This commutation follows from (2.257) and

$$\underline{\mathbf{F}}_\alpha(\vartheta_{\mathrm{a}}(\lambda)f) = |\lambda|\,\vartheta_{\mathrm{a}}(\lambda^{-1})\underline{\mathbf{F}}_\alpha(f)\,, \quad \forall f \in L^2(X_{\mathbb{Q},S})\,. \tag{2.261}$$

Thus there exists a unitary function z on $\widehat{C_{\mathbb{Q},S}}$ such that

$$I \circ w \circ \underline{\mathbf{F}}_\alpha \circ w^{-1} = \mathbf{F}_C^{-1} \circ z \circ \mathbf{F}_C$$

and it remains to show that $z = u$. To get this we let $\chi \in \widehat{C_{\mathbb{Q},S}}$ be a character of $C_{\mathbb{Q},S}$ and consider the distribution on $\mathbb{A}_{\mathbb{Q},S}$

$$\Delta_\chi(f) = \int_{C_{\mathbb{Q},S}} \mathfrak{E}(f)(x)\chi(x) d^*x\,, \quad \forall f \in \mathcal{S}(\mathbb{A}_{\mathbb{Q},S})\,. \tag{2.262}$$

This is well defined by the exponential decay of $\mathfrak{E}(f)(x)$ and gives a homogeneous distribution. The computation of the Fourier transform of homogeneous distributions on $\mathbb{A}_{\mathbb{Q},S}$ thus gives, using (2.258), the equality

$$\int_{C_{\mathbb{Q},S}} \mathfrak{E}(\mathbf{F}_\alpha(f))(x^{-1})\chi(x)d^*x = \prod u_v(\chi_v) \int_{C_{\mathbb{Q},S}} \mathfrak{E}(f)(x)\chi(x)d^*x \tag{2.263}$$

where $\tilde{\chi} = \prod \chi_v$ is the decomposition of the composition $\tilde{\chi}$ of χ with the quotient map $\mathbb{A}^*_{\mathbb{Q},S} \to C_{\mathbb{Q},S}$. Since $\chi_v = \pi_v(\chi)$ this shows that $z = u$. □

Now we proceed as in Section 3 and we introduce infrared and ultraviolet cutoffs. For the infrared one we use the orthogonal projection P_Λ onto the subspace

$$\{\xi \in L^2(X_S) \,|\, \xi(x) = 0, \forall x \text{ with } |x| > \Lambda\}. \tag{2.264}$$

Thus, P_Λ is the operator acting as multiplication by the function ρ_Λ, with $\rho_\Lambda(x) = 1$ for $|x| \le \Lambda$ and $\rho_\Lambda(x) = 0$ for $|x| > \Lambda$. This gives an infrared cutoff.
To get an ultraviolet cutoff we use $\widehat{P}_\Lambda = \underline{\mathbf{F}}_\alpha P_\Lambda \underline{\mathbf{F}}_\alpha^{-1}$, where $\underline{\mathbf{F}}_\alpha$ is the Fourier transform of Lemma 2.33, which depends upon the choice of the basic character $\alpha = \prod_v \alpha_v$. Finally, we let

$$R_\Lambda = \widehat{P}_\Lambda P_\Lambda. \tag{2.265}$$

We shall now prove an analog of Lemma 2.21 in this context. We first need to extend the quantized calculus of Definition 2.19 to arbitrary modulated groups C (Definition 2.8).
We let C be such a group and we let $\widehat{C}$ denote the Pontrjagin dual of C. The elements of $L^\infty(\widehat{C})$ act as multiplication operators on the Hilbert space $\mathcal{H} := L^2(\widehat{C})$. We then consider an operator H on $\mathcal{H}$ of the form

$$H := 2\mathbf{F}_C\, \mathbf{1}_P\, \mathbf{F}_C^{-1} - 1, \tag{2.266}$$

where $\mathbf{F}_C : L^2(C) \to \mathcal{H}$ is the Fourier transform and $\mathbf{1}_P$ is the multiplication by the characteristic function of the set $P = \{u \in C \,|\, |u| \ge 1\}$. Here we use the notation $\mathrm{Mod}(u) = |u|$ for $u \in C$. The operator (2.266) plays the role of the Hilbert transform.
As in (2.137) above, we define the "quantized" differential of f as the operator

$$\text{đ}f := [H,\, f] = Hf - fH. \tag{2.267}$$

We introduce the notation

$$\Pi_{[a,b]} := \mathbf{F}_C\, \mathbf{1}_Y\, \mathbf{F}_C^{-1}\,, \quad \text{for } Y = \{u \in C \,|\, e^{2\pi a} \le |u| \le e^{2\pi b}\}, \tag{2.268}$$

whose normalization fits with (2.136). We then have the following result.

Lemma 2.35. *Let S be as above. For any Λ there is a unitary operator*

$$W = W_\Lambda : L^2(X_{\mathbb{Q},S}) \to L^2(\widehat{C_{\mathbb{Q},S}}) \tag{2.269}$$

such that, for any $h_j \in \mathcal{S}(C_{\mathbb{Q},S})$, $j = 1, 2$, one has

$$\begin{aligned} &W\, \vartheta_{\rm a}(\tilde{h}_1)\widehat{P}_\Lambda P_\Lambda \vartheta_{\rm a}(\tilde{h}_2)W^* = \\ &\hat{h}_1 \left(\tfrac{1}{2}u^{-1}\text{đ}u\, \Pi_{[-\infty, \frac{2\log\Lambda}{2\pi}]} + \Pi_{[0, \frac{2\log\Lambda}{2\pi}]}\right)\hat{h}_2. \end{aligned} \tag{2.270}$$

Here $\tilde{h}_j(\lambda) = |\lambda|^{-1/2} h_j(\lambda)$, *the operator* $đu$ *is the quantized differential of the function* $u := \prod_S u_v \circ \pi_v$ *of* (2.260) *and* $\hat{h}_j$ *is the multiplication operator by the Fourier transform* $\mathbf{F}_C(h_j)$.

PROOF. The proof is the same as for Lemma 2.21 using Lemma 2.34 instead of Lemma 2.20 and the identity (2.257) instead of (2.157). ☐

We now give a different proof of the following result of [**71**]. This proof is based on the idea of [**37**] and was explained in the class [**72**] of 1999.

THEOREM 2.36. *Let* $\mathbb{A}_{\mathbb{Q},S}$ *be as above, with basic character* $\alpha = \prod_{v \in S} \alpha_v$. *Let* $h \in \mathcal{S}(C_{\mathbb{Q},S})$ *be a function with compact support. Then, in the limit* $\Lambda \to \infty$, *one has*

$$\operatorname{Tr}(\vartheta_{\mathrm{a}}(h) R_\Lambda) = 2h(1)\log\Lambda + \sum_{v \in S} \int'_{\mathbb{Q}_v^*} \frac{h(u^{-1})}{|1-u|}\, d^*u + o(1). \tag{2.271}$$

Here each $\mathbb{Q}_v^*$ is embedded in $C_{\mathbb{Q},S}$ by the map $u \mapsto (1, 1, \ldots, u, \ldots, 1)$. The principal value $\int'$ is as in Definition 2.12.

PROOF. Let us first show that $\vartheta_{\mathrm{a}}(h) R_\Lambda$ is a trace-class operator. We use (2.270) and treat separately the two terms. The operator $\hat{h}\, \Pi_{[0, \frac{2\log\Lambda}{2\pi}]}$ is of trace-class and its trace gives the term $2h(1)\log\Lambda$. To show that the operator $\hat{h}\, \frac{1}{2} u^{-1} đu\, \Pi_{[-\infty, \frac{2\log\Lambda}{2\pi}]}$ is of trace-class, it is enough to show that $\hat{h}\, \frac{1}{2} u^{-1} đu$ is of trace class. Moreover, since the increasing family of projections $\Pi_{[-\infty, \frac{2\log\Lambda}{2\pi}]} \to 1$ when $\Lambda \to \infty$, one then gets the convergence of the traces

$$\operatorname{Tr}(\hat{h}\, \frac{1}{2} u^{-1} đu\, \Pi_{[-\infty, \frac{2\log\Lambda}{2\pi}]}) \to \operatorname{Tr}(\hat{h}\, \frac{1}{2} u^{-1} đu)\,. \tag{2.272}$$

Next one has $u^{-1} đu = -đu^{-1} u$ and $\hat{h} đu^{-1} u = đ(\hat{h} u^{-1}) u - đ\hat{h}$. By [**68**] Chapter 4, the operator $đf$ is of infinite order and hence of trace-class for any Schwartz function f. This is the case for $\hat{h}$ and to show that this also holds for $\hat{h} u^{-1}$ we need to show that all the derivatives of the function u are of tempered growth. This is done separately for the u_v. When v is a non-Archimedean place u_v is a periodic smooth function so the answer follows. When $v = \infty$ is the Archimedean place, one uses the formula (with γ the Euler constant)

$$\frac{\Gamma'}{\Gamma}(z) = -\gamma + \sum_0^\infty \left(\frac{1}{n+1} - \frac{1}{n+z}\right)$$

together with (2.153) to get an estimate of the form $u_\infty^{(n)}(s) = O(\log|s|^n)$ for $|s| \to \infty$. The trace of $\vartheta_{\mathrm{a}}(h) R_\Lambda$ can then be computed in two different ways. The first uses the formula

$$\operatorname{Tr}(\hat{h}\, u^{-1} đu) = \sum_{v \in S} \operatorname{Tr}(\hat{h}\, (u_v \circ \pi_v)^{-1} đ(u_v \circ \pi_v)) \tag{2.273}$$

which follows from (2.260), the derivation property of the quantized differential, and the cyclic property of the trace.

The second way is direct and geometric ([**71**]) and gives (2.271) with the prescription of Definition 2.12 for the principal values. It thus remains to show that the two results are the same. This follows from the reformulation by Weil [**294**] of the explicit formulas of number theory and will be explained in §8. ☐

8. Weil's formulation of the explicit formulas

In this section we shall recall Weil's formulation of Riemann's explicit formulas in number theory. It will prepare the ground for the more complicated L-functions treated in the §10. We shall also check by explicit computations that the rewriting of the Weil explicit formulas which is predicted by the trace formula of Theorem 2.36 is correct. Our computation is straightforward and it will be useful to obtain the principal values as explicitly as possible, but the result can be better understood conceptually using [**163**]. We shall also show that the Weil principal values can be computed also in the complex case using the minimal subtraction scheme $\overline{\mathrm{MS}}_K$ which we already met in §4.2 for $K = \mathbb{R}$.

8.1. L-functions.

One lets $\mathbb{K}$ be a global field. It is by definition a field which embeds as a discrete cocompact subfield in a (non-discrete) locally compact semisimple abelian ring A. The topological ring A then depends functorially on $\mathbb{K}$ and is called the Adèle ring $\mathbb{A}_{\mathbb{K}}$ of $\mathbb{K}$. Thus, though the field $\mathbb{K}$ itself has no interesting topology, there is a canonical and highly nontrivial topological ring which is canonically associated to $\mathbb{K}$.

In general for a non-Archimedean local field K we use the notations $\mathcal{O}$ for the maximal compact subring, $\mathfrak{m}$ for the maximal ideal of $\mathcal{O}$, π for a generator of the ideal $\mathfrak{m}$ (i.e., $\mathfrak{m} = \pi\mathcal{O}$).

Let $\mathbb{K}$ be a global field and $\mathbb{A}_{\mathbb{K}}$ the ring of adèles of $\mathbb{K}$. It is the restricted product (Definition 2.25) of the local fields $\mathbb{K}_v$ indexed by the set of places v of $\mathbb{K}$, with respect to the maximal compact subrings $\mathcal{O}_v$. Similarly, the Bruhat-Schwartz space $\mathcal{S}(\mathbb{A}_{\mathbb{K}})$ is the restricted tensor product of the local Bruhat-Schwartz spaces $\mathcal{S}(\mathbb{K}_v)$, with respect to the vectors $\mathbf{1}_{\mathcal{O}_v}$.

L-functions on $\mathbb{K}$ are associated to Grössencharakters, i.e., to characters of the idèles class group

$$C_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^* \,, \tag{2.274}$$

where $\mathbb{A}_{\mathbb{K}}^* = \mathrm{GL}_1(\mathbb{A}_{\mathbb{K}})$ is the group of idèles. Let χ be a character of the idèles class group, we consider χ as a character of $\mathbb{A}_{\mathbb{K}}^*$ which is 1 on $\mathbb{K}^*$. As such it can be written as a product

$$\chi(j) = \prod \chi_v(j_v)\,, \quad j = (j_v) \in \mathbb{A}_{\mathbb{K}}^* \,. \tag{2.275}$$

By considering the restriction of χ to the compact subgroup

$$G_0 = \prod \mathcal{O}_v^* \times 1 \subset \mathbb{A}_{\mathbb{K}}^* \,, \tag{2.276}$$

it follows that for all finite v but a finite number, one has

$$\chi_v|_{\mathcal{O}_v^*} = 1 \,. \tag{2.277}$$

One says that χ is unramified at v when this holds. Then $\chi_v(x)$ only depends upon the module $|x|$, since

$$\mathbb{K}_v^*/\mathcal{O}_v^* = \mathrm{Mod}(\mathbb{K}_v) \,. \tag{2.278}$$

Thus χ_v is determined by

$$\chi_v(\pi_v) \tag{2.279}$$

which does not depend upon the choice of π_v ($\mathrm{Mod}\,\mathcal{O}_v^*$).

Let χ be a quasi-character of $C_{\mathbb{K}}$, of the form

$$\chi(x) = \chi_0(x)\,|x|^s \tag{2.280}$$

where $s \in \mathbb{C}$ and χ_0 is a character of $C_{\mathbb{K}}$. The real part σ of s is uniquely determined by

$$|\chi(x)| = |x|^\sigma\,. \tag{2.281}$$

Let P be the finite set of finite places where χ_0 is ramified. The L-function $L(\chi_0, s)$ is defined for $\sigma = \Re(s) > 1$ as

$$L(\chi_0, s) = \prod_{\substack{v \text{ finite} \\ v \notin P}} (1 - \chi_{0,v}(\pi_v)\, q_v^{-s})^{-1} = \prod_{\substack{v \text{ finite} \\ v \notin P}} (1 - \chi_v(\pi_v))^{-1} \tag{2.282}$$

where

$$|\pi_v| = q_v^{-1}. \tag{2.283}$$

8.2. Weil's explicit formula.

Let us recall the Weil explicit formulas [**294**]. Let as above $\mathbb{K}$ be a global field. The idèle class group $C_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$ is a modulated group, where the module Mod is defined (cf. [**293**]) using the scaling of the additive Haar measure of $\mathbb{A}_{\mathbb{K}}$ under the action of $\mathbb{A}_{\mathbb{K}}^*$ by multiplication. One lets $C_{\mathbb{K},1} \subset C_{\mathbb{K}}$ be the kernel of the module. One identifies the quotient $C_{\mathbb{K}}/C_{\mathbb{K},1}$ with the range of the module

$$N = \{|g| \,|\, g \in C_{\mathbb{K}}\} \subset \mathbb{R}_+^*\,. \tag{2.284}$$

One endows N with its normalized Haar measure d^*x. Given a function F on N such that, for some $b > \frac{1}{2}$,

$$|F(\nu)| = O(\nu^b) \quad \nu \to 0\ ,\ |F(\nu)| = O(\nu^{-b})\ ,\ \nu \to \infty\,, \tag{2.285}$$

one lets

$$\Phi(s) = \int_N F(\nu)\, \nu^{1/2 - s}\, d^*\nu\,. \tag{2.286}$$

Given a Grössencharakter χ, i.e., a character of $C_{\mathbb{K}}$ and any ρ in the strip $0 < \Re(\rho) < 1$, one lets $N(\chi, \rho)$ be the order of vanishing of $L(\chi, s)$ at $s = \rho$. One lets

$$S(\chi, F) = \sum_{\rho \in \mathbb{C}/N^\perp} N(\chi, \rho)\, \Phi(\rho)\,, \tag{2.287}$$

where the sum runs over ρ's in the above open strip, modulo the subgroup $N^\perp := \{s \in \mathbb{C} |\, \nu^s = 1,\ \forall \nu \in N\}$. One then defines a distribution $\mathbf{D}_{\mathbb{K}}$ on $C_{\mathbb{K}}$ by

$$\mathbf{D}_{\mathbb{K}} = \log |d^{-1}|\, \delta_1 + \mathbf{D} - \sum_v \mathbf{D}_v\,, \tag{2.288}$$

where δ_1 is the Dirac mass at $1 \in C_{\mathbb{K}}$, where d is a differental idèle of $\mathbb{K}$ so that $|d|^{-1}$ is up to sign the discriminant of $\mathbb{K}$ when $\mathrm{char}(\mathbb{K}) = 0$, and is q^{2g-2} when $\mathbb{K}$ is a function field over a curve of genus g with coefficients in the finite field $\mathbb{F}_q$.

The distribution $\mathbf{D}$ is given by

$$\mathbf{D}(f) = \int_{C_{\mathbb{K}}} f(w)\,(|w|^{1/2} + |w|^{-1/2})\, d^*w\,, \tag{2.289}$$

where the Haar measure d^*w is normalized. The distributions $\mathbf{D}_v$ are parameterized by the places v of $\mathbb{K}$ and are obtained as follows. For each v one considers the natural proper homomorphism

$$\mathbb{K}_v^* \to C_{\mathbb{K}}\ ,\ x \mapsto \text{class of } (1, \ldots, x, 1, \ldots) \tag{2.290}$$

of the multiplicative group of the local field $\mathbb{K}_v$ to the idèle class group $C_{\mathbb{K}}$.
One then has

$$\mathbf{D}_v(f) = Pfw \int_{\mathbb{K}_v^*} \frac{f(u)}{|1-u|}\, |u|^{1/2}\, d^*u\,, \tag{2.291}$$

where the Haar measure d^*u is normalized, and where the Weil principal value Pfw of the integral is defined, for a local field $K = \mathbb{K}_v$, as

$$Pfw \int_{\mathbb{K}_v^*} \mathbf{1}_{\mathcal{O}_v^*} \frac{1}{|1-u|}\, d^*u = 0\,, \tag{2.292}$$

if the local field $\mathbb{K}_v$ is non-Archimedean, and otherwise as

$$Pfw \int_{\mathbb{K}_v^*} \varphi(u)\, d^*u = PF_0 \int_{\mathbb{R}_+^*} \psi(\nu)\, d^*\nu\,. \tag{2.293}$$

Here $\psi(\nu) = \int_{|u|=\nu} \varphi(u) d_\nu u$ is obtained by integrating φ over the fibers, while

$$\begin{aligned} &PF_0 \textstyle\int \psi(\nu) d^*\nu = \\ &\qquad 2\log(2\pi)c + \lim_{t\to\infty} \left(\textstyle\int (1 - f_0^{2t})\psi(\nu) d^*\nu - 2c\log t\right)\,. \end{aligned} \tag{2.294}$$

where one assumes that $\psi - c\ f_1^{-1}$ is integrable on $\mathbb{R}_+^*$, and

$$f_0(\nu) = \inf(\nu^{1/2}, \nu^{-1/2}) \qquad \forall\, \nu \in \mathbb{R}_+^*, \qquad f_1 = f_0^{-1} - f_0\,. \tag{2.295}$$

The Weil explicit formula is the content of the following statement.

THEOREM 2.37. *([**294**]) With the above notations one has*

$$S(\chi, F) = \mathbf{D}_{\mathbb{K}}(F(|w|)\,\chi(w))\,. \tag{2.296}$$

8.3. Fourier transform on $C_{\mathbb{K}}$.

We have a non-canonical isomorphism

$$C_{\mathbb{K}} \simeq C_{\mathbb{K},1} \times N, \tag{2.297}$$

where $N \subset \mathbb{R}_+^*$ is the range of the norm $|\cdot| : C_{\mathbb{K}} \to \mathbb{R}_+^*$. For number fields this is $N = \mathbb{R}_+^*$, while for function fields in positive characteristic $N \simeq \mathbb{Z}$ is the subgroup $q^{\mathbb{Z}} \subset \mathbb{R}_+^*$ with $q = p^\ell$ the cardinality of the field of constants.

We denote by $\widehat{C_{\mathbb{K},1}}$ the group of characters of the compact subgroup $C_{\mathbb{K},1} \subset C_{\mathbb{K}}$, i.e. the Pontrjagin dual of $C_{\mathbb{K},1}$. Given a character χ of $C_{\mathbb{K},1}$, we let $\tilde{\chi}$ denote the unique extension of χ to $C_{\mathbb{K}}$ which is equal to one on N.

For $f \in \mathcal{S}(C_{\mathbb{K}})$, with compact support, and $\tilde{\chi}$ the extension of a character $\chi \in \widehat{C_{\mathbb{K},1}}$ to $C_{\mathbb{K}}$ as above, we consider the Fourier transform

$$\hat{f}(\tilde{\chi}, \rho) = \int_{C_{\mathbb{K}}} f(u)\tilde{\chi}(u)\,|u|^{\rho}\,d^*u. \tag{2.298}$$

DEFINITION 2.38. *Consider the bicharacter*

$$\langle z, \lambda\rangle := \lambda^z, \qquad \text{for } (z,\lambda) \in \mathbb{C} \times \mathbb{R}_+^*. \tag{2.299}$$

Let N denote as above the range of the norm $|\cdot| : C_{\mathbb{K}} \to \mathbb{R}_+^*$. Then $N^{\perp} \subset \mathbb{C}$ denotes the subgroup

$$N^{\perp} := \{z \in \mathbb{C} |\, \lambda^z = 1,\ \forall \lambda \in N\}. \tag{2.300}$$

In the case where $N = q^{\mathbb{Z}}$ (function fields), the subgroup $N^{\perp}$ is nontrivial and given by

$$N^{\perp} = \frac{2\pi i}{\log q}\mathbb{Z}. \tag{2.301}$$

Since in the function field case the L-functions are functions of q^{-s}, there is a periodicity by $N^{\perp}$, hence we need to consider $\rho \in \mathbb{C}/N^{\perp}$ in (2.287). In the number field case this does not matter, since $N = \mathbb{R}_+^*$ and $N^{\perp}$ is trivial.

8.4. Computation of the principal values.

We now compare the principal values Pfw of (2.296) with those of Theorem 2.36. We fix a decomposition $C_{\mathbb{K}} = C_{\mathbb{K},1} \times N$ as in (2.297). We let χ_0 be a character of $C_{\mathbb{K},1}$ and let $\tilde{\chi}_0$ denote the unique extension of χ_0 to $C_{\mathbb{K}}$ which is equal to one on N. Let h be the function on $C_{\mathbb{K}}$ given by

$$|g|^{-1/2}\,h(g^{-1}) = F(|g|)\,\tilde{\chi}_0(g)\,. \tag{2.302}$$

We rewrite the equality (2.286) in terms of h. By (2.286) one has

$$\Phi\left(\frac{1}{2} + is\right) = \int_{C_{\mathbb{K}}} F(|g|)\,|g|^{-is}\,d^*g\,. \tag{2.303}$$

Thus, in terms of h,

$$\int h(g)\,\tilde{\chi}_0(g)\,|g|^{1/2+is}\,d^*g = \Phi\left(\frac{1}{2} + is\right)\,. \tag{2.304}$$

while if χ_1 is a character of $C_{\mathbb{K},1}$ with $\chi_1 \neq \chi_0$, one gets

$$\int h(g)\,\tilde{\chi}_1(g)\,|g|^{1/2+is}\,d^*g = \int F(|g^{-1}|)\,\tilde{\chi}_0(g^{-1})\,\tilde{\chi}_1(g)\,|g|^{is}\,d^*g \,=\, 0\,.$$

Thus, with our notations (2.298) we see that

$$\operatorname{Supp}\hat{h} \subset \chi_0 \times \mathbb{R}\ ,\ \hat{h}(\tilde{\chi}_0, \rho) = \Phi(\rho)\,. \tag{2.305}$$

Thus we can write, using the notation of a set with multiplicity for the set $Z_{\tilde{\chi}}$ of zeros of $L(\tilde{\chi}, \rho)$ in the strip $0 < \Re\,\rho < 1$ modulo $N^{\perp}$,

$$S(\tilde{\chi}_0, F) = \sum_{\chi \in \widehat{C_{\mathbb{K},1}}} \sum_{Z_{\tilde{\chi}}} \hat{h}(\tilde{\chi}, \rho)\,. \tag{2.306}$$

Let us now evaluate each term in (2.288). The first gives $(\log|d^{-1}|)\,h(1)$. One has, using (2.289) and (2.302),

$$\begin{aligned}\langle \mathbf{D}, F(|g|)\,\tilde{\chi}_0(g)\rangle &= \int_{C_{\mathbb{K}}} |g|^{-1/2}\,h(g^{-1})\,(|g|^{1/2}+|g|^{-1/2})\,d^*g \\ &= \int_{C_{\mathbb{K}}} h(u)\,(1+|u|)\,d^*u = \hat{h}(0)+\hat{h}(1)\end{aligned}$$

where for the trivial character of $C_{\mathbb{K},1}$ one uses the notation

$$\hat{h}(z) = \hat{h}(1,z) \qquad \forall\, z \in \mathbb{C}\,. \tag{2.307}$$

Thus the first two terms of (2.288) give

$$(\log|d^{-1}|)\,h(1) + \hat{h}(0) + \hat{h}(1)\,. \tag{2.308}$$

Let then v be a place of $\mathbb{K}$. One has by (2.291) and (2.302),

$$\langle \mathbf{D}_v, F(|g|)\,\tilde{\chi}_0(g)\rangle = Pfw \int_{\mathbb{K}_v^*} h(u^{-1})|1-u|\,d^*u\,. \tag{2.309}$$

We can thus write the contribution of the last term of (2.288) as

$$-\sum_v Pfw \int_{\mathbb{K}_v^*} h(u^{-1})|1-u|\,d^*u\,. \tag{2.310}$$

Thus the equality of Weil can be rewritten as

$$\begin{aligned}&\hat{h}(0)+\hat{h}(1) - \textstyle\sum_{\chi\in\widehat{C_{\mathbb{K},1}}} \sum_{Z_{\tilde{\chi}}} \hat{h}(\tilde{\chi},\rho) = \\ &\qquad (\log|d|)\,h(1) + \textstyle\sum_v Pfw \int_{\mathbb{K}_v^*} \frac{h(u^{-1})}{|1-u|}\,d^*u\,,\end{aligned} \tag{2.311}$$

which holds for finite linear combinations of functions h of the form (2.302).

Let us now compare the Weil principal values with those dictated by Theorem 2.36. We first work with a local field K and compare (2.292), (2.293) with our prescription.

8.4.1. *Non-Archimedean case.*

Let first K be non-Archimedean. Let α be a character of K such that

$$\alpha|_{\mathcal{O}} = 1\ ,\ \ \alpha|_{\pi^{-1}\mathcal{O}} \neq 1\,. \tag{2.312}$$

Then, for the Fourier transform given by

$$(\mathbf{F}_\alpha f)(x) = \int f(y)\,\alpha(xy)\,dy\,, \tag{2.313}$$

with dy the selfdual Haar measure, one has

$$\mathbf{F}_\alpha(\mathbf{1}_{\mathcal{O}}) = \mathbf{1}_{\mathcal{O}}\,. \tag{2.314}$$

LEMMA 2.39. *With the above choice of α one has*

$$\int' \frac{h(u^{-1})}{|1-u|}\,d^*u = Pfw \int \frac{h(u^{-1})}{|1-u|}\,d^*u \tag{2.315}$$

with the notations of Theorem 2.36.

PROOF. By construction the two sides can only differ by a multiple of $h(1)$. Let us recall from Theorem 2.36 that the left hand side is given by

$$\left\langle L, \frac{h(u^{-1})}{|u|} \right\rangle , \tag{2.316}$$

where L is the unique extension of $\rho^{-1}\frac{du}{|1-u|}$ whose Fourier transform vanishes at 1, $\hat{L}(1) = 0$. Thus from (2.292) we just need to check that (2.316) vanishes for $h = \mathbf{1}_{\mathcal{O}^*}$, i.e., that

$$\langle L, \mathbf{1}_{\mathcal{O}^*} \rangle = 0 . \tag{2.317}$$

Equivalently, shifting both sides by $1 \in K$ and using the Parseval formula (2.96), we need to show that

$$\langle \log |u|, \hat{\mathbf{1}}_Y \rangle = 0 . \tag{2.318}$$

where $Y = \{y \in K \,;\, |y-1| = 1\}$ so that $\mathbf{1}_Y(y) = \mathbf{1}_{\mathcal{O}^*}(y-1)$. One has

$$\hat{\mathbf{1}}_Y(x) = \int_Y \alpha(xy)\, dy = \int_{\mathcal{O}^*} \alpha(x(1+z))\, dz = \alpha(x)\, \hat{\mathbf{1}}_{\mathcal{O}^*}(x)$$

and $\mathbf{1}_{\mathcal{O}^*} = \mathbf{1}_{\mathcal{O}} - \mathbf{1}_{\mathfrak{m}}$, $\hat{\mathbf{1}}_{\mathcal{O}^*} = \mathbf{1}_{\mathcal{O}} - |\pi|\, \mathbf{1}_{\pi^{-1}\mathcal{O}}$, thus, with $q^{-1} = |\pi|$,

$$\hat{\mathbf{1}}_Y(x) = \alpha(x) \left(\mathbf{1}_{\mathcal{O}} - \frac{1}{q}\, \mathbf{1}_{\pi^{-1}\mathcal{O}} \right)(x) . \tag{2.319}$$

We now need to compute $\int \log |x|\, \hat{\mathbf{1}}_Y(x)\, dx = A + B$,

$$A = -\frac{1}{q} \int_{\pi^{-1}\mathcal{O}^*} \alpha(x)\,(\log q)\, dx \;,\; B = \left(1 - \frac{1}{q}\right) \int_{\mathcal{O}} \log |x|\, dx . \tag{2.320}$$

Let us show that $A + B = 0$. One has $\int_{\mathcal{O}} dx = 1$, and

$$A = \; - \int_{\mathcal{O}^*} \alpha(\pi^{-1}y)\,(\log q)\, dy = -\log q \left(\int_{\mathcal{O}} \alpha(\pi^{-1}y)\, dy - \int_{\mathfrak{m}} dy \right) = \; \frac{1}{q}\, \log q$$

since $\int_{\mathcal{O}} \alpha(\pi^{-1}y)\, dy = 0$ as $\hat{\mathbf{1}}_{\mathcal{O}}(\pi^{-1}) = 0$.

To compute B, note that $\int_{\pi^n\mathcal{O}^*} dy = q^{-n}\left(1 - \frac{1}{q}\right)$ so that

$$B = \left(1 - \frac{1}{q}\right)^2 \sum_{n=0}^{\infty} (-n \log q)\, q^{-n} = -q^{-1} \log q . \tag{2.321}$$

and $A + B = 0$. ∎

8.4.2. *Real case.*

Let us now treat the case of Archimedean fields. We take $K = \mathbb{R}$ first, and we choose the additive character $e_{\mathbb{R}}(x) = e^{-2\pi i x}$.

PROPOSITION 2.40. *With the notations of Theorem 2.36, one has*

$$\int_{\mathbb{R}}' f(u)\, \frac{1}{|1-u|}\, d^*u = Pfw \int_{\mathbb{R}} f(u)\, \frac{1}{|1-u|}\, d^*u . \tag{2.322}$$

PROOF. With the notations of (2.293), and f_0, f_1 as in (2.295), one has

$$Pfw \int_{\mathbb{R}^*} f_0^3(|u|)\, \frac{|u|^{1/2}}{|1-u|}\, d^*u = \log\pi + \gamma \tag{2.323}$$

with $\gamma = -\Gamma'(1)$. Indeed, integrating over the fibers gives $\psi = f_0^4 \times (1-f_0^4)^{-1}$, and one gets

$$PF_0 \int_{\mathbb{R}_+^*} \psi\, d^*u = \log 2\pi + \lim_{t\to\infty}\left(\int_{\mathbb{R}_+^*} (1-f_0^{2t})\, f_0^4(1-f_0^4)^{-1}\, d^*u - \log t\right)$$
$$= \log 2\pi + \lim_{t\to\infty}\left(\int_0^1 \frac{1-v^{t/2}}{1-v}\, dv - \log t\right) = \log 2\pi + \gamma - \log 2$$

Taking $f(u) = |u|^{1/2}\, f_0^3(|u|)$, the right hand side of (2.107) gives $\lambda - \log 2 = \log\pi + \gamma$. Indeed, one has $f(1) = 1$ and

$$\lim_{\varepsilon\to 0}\left(\int_{|1-u|\geq\varepsilon} f_0^3(|u|)\, \frac{|u|^{1/2}}{|1-u|}\, d^*u + \log\varepsilon\right)$$
$$= 2\lim_{\varepsilon\to 0}\left(\int_{|1-u|\geq\varepsilon, |u|<1} f_0^3(|u|)\, \frac{|u|^{1/2}}{|1-u|}\, d^*u + \frac{1}{2}\log\varepsilon\right)$$
$$= \lim_{\varepsilon\to 0}\left(\int_0^{1-\varepsilon} \frac{2udu}{1-u^2} + \log\varepsilon\right) = -\log 2\,.$$

Thus we conclude, using (2.323), that (2.322) holds. □

8.4.3. *Complex case.*

Let us finally consider the case $K = \mathbb{C}$. We choose the basic character α as

$$e_{\mathbb{C}}(z) = \exp -2\pi i(z+\bar{z})\,; \tag{2.324}$$

the selfdual Haar measure is $dz\, d\bar{z} = |dz \wedge d\bar{z}| = d\rho^2 d\theta$ for $z = \rho\, e^{i\theta}$, and the function $f(z) = \exp -2\pi|z|^2$ is selfdual.

The normalized multiplicative Haar measure is

$$d^*z = \frac{|dz\wedge d\bar{z}|}{2\pi|z|^2} = \frac{d\rho^2}{\rho^2}\frac{d\theta}{2\pi}\,, \quad z = \rho\, e^{i\theta}\,. \tag{2.325}$$

PROPOSITION 2.41.

(1) *The Fourier transform* $\hat{\varphi} = \mathbf{F}_{e_{\mathbb{C}}}(\varphi)$ *of* $\varphi(z) = -\log|z|_{\mathbb{C}}$ *is given by the distribution*

$$L(f) := 2(\log 2\pi + \gamma) f(0) + \lim_{\varepsilon\to 0}\left(\int_{|u|_{\mathbb{C}}\geq\varepsilon} f(u)\, d^*u + \log\varepsilon\, f(0)\right). \tag{2.326}$$

(2) *Let* $\lambda' = 2(\log 2\pi + \gamma)$*; Then one has for any test function* f*,*

$$\begin{aligned} &\textstyle\int_{\mathbb{C}}' f(u)\, \frac{1}{|1-u|_{\mathbb{C}}}\, d^*u = \lambda'\, f(1) + \lim_{\varepsilon\to 0}\left(\int \frac{f(u)}{|1-u|_{\mathbb{C}}^{1-\varepsilon}}\, d^*u - \frac{1}{\varepsilon}\, f(1)\right) \\ &\quad = \lambda'\, f(1) + \lim_{\varepsilon\to 0}\left(\int_{|1-u|_{\mathbb{C}}\geq\varepsilon} \frac{f(u)}{|1-u|_{\mathbb{C}}}\, d^*u + (\log\varepsilon) f(1)\right)\,. \end{aligned} \tag{2.327}$$

(3) *With the notations of Theorem 2.36, one has*

$$\int'_{\mathbb{C}} f(u)\,\frac{1}{|1-u|_{\mathbb{C}}}\,d^*u = Pfw \int_{\mathbb{C}} f(u)\,\frac{1}{|1-u|_{\mathbb{C}}}\,d^*u\,. \tag{2.328}$$

PROOF. (1) Let us compute the Fourier transform $\hat{\varphi}$ of the distribution

$$\varphi(z) = -\log|z|_{\mathbb{C}} = -2\log|z|\,. \tag{2.329}$$

One has

$$\langle \varphi, \exp(-2\pi|z|^2)\rangle = \log 2\pi + \gamma\,, \tag{2.330}$$

as is seen using $\frac{\partial}{\partial\varepsilon}\left(\int e^{-2\pi|z|^2}\,|z|^{-2\varepsilon}\,|dz\wedge d\bar{z}|\right) = \frac{\partial}{\partial\varepsilon}\left((2\pi)^{\varepsilon}\,\Gamma(1-\varepsilon)\right)$.
Thus $\langle\hat{\varphi}, \exp -2\pi|u|^2\rangle = \log 2\pi + \gamma$ and one gets

$$\langle\hat{\varphi}, f\rangle = \lim_{\varepsilon\to 0}\left(\int_{|u|_{\mathbb{C}}\geq\varepsilon} f(u)\,d^*u + \log\varepsilon\, f(0)\right) + \lambda'\, f(0) \tag{2.331}$$

To see this one uses the analogue of (2.111) for $K = \mathbb{C}$, i.e.

$$\begin{aligned} &\lim_{\varepsilon\to 0}\left(\int_{|u|_{\mathbb{C}}\geq\varepsilon} f(u)\,d^*u + \log\varepsilon\, f(0)\right) = \\ &\qquad \lim_{\varepsilon\to 0}\left(\int f(x)\,|x|_{\mathbb{C}}^{\varepsilon}\,d^*x - \tfrac{1}{\varepsilon}\,f(0)\right), \end{aligned} \tag{2.332}$$

to compute the right hand side of (2.331) for $f(z) = \exp(-2\pi|z|^2)$. The first term gives $-\log 2\pi - \gamma$ and after adding $\lambda' f(0)$ one gets the same as (2.330). Thus we get (2.327) and (2).

(3) Let us compare this principal value with Pfw. When one integrates over the fibers of $|\ .\ |_{\mathbb{C}}\ :\ \mathbb{C}^* \longrightarrow \mathbb{R}_+^*$ the function $|1-z|_{\mathbb{C}}^{-1}$, one gets

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{|1-e^{i\theta}z|^2}\,d\theta = \frac{1}{1-|z|^2}\ \text{ if } |z|<1\,,\ \text{ and }\ \frac{1}{|z|^2-1}\ \text{ if } |z|>1\,. \tag{2.333}$$

Thus for any test function f on $\mathbb{R}_+^*$ one has, by (2.293),

$$Pfw\int f(|u|_{\mathbb{C}})\,\frac{1}{|1-u|_{\mathbb{C}}}\,d^*u = PF_0\int f(\nu)\,\frac{1}{|1-\nu|}\,d^*\nu \tag{2.334}$$

with the notations of (2.294). With f_0, f_1 as in (2.295) and $f_2(\nu) = \nu^{\frac{1}{2}}\,f_0(\nu)$ we thus get, using (2.294),

$$Pfw\int f_2(|u|_{\mathbb{C}})\,\frac{1}{|1-u|_{\mathbb{C}}}\,d^*u = PF_0\int f_0\,f_1^{-1}d^*\nu = 2(\log 2\pi + \gamma)\,. \tag{2.335}$$

since with $\psi = f_0\,f_1^{-1}$ one has $c = 1$ and

$$\lim_{t\to\infty}\left(\int (1-f_0^{2t})\,\psi(\nu)\,d^*\nu - 2c\log t\right) = 2\lim_{t\to\infty}\left(\int_0^1\frac{1-\nu^t}{1-\nu}d\nu - \log t\right) = 2\gamma\,.$$

We shall now show that

$$\lim_{\varepsilon\to 0}\left(\int_{|1-u|_{\mathbb{C}}\geq\varepsilon}\frac{f_2(|u|_{\mathbb{C}})}{|1-u|_{\mathbb{C}}}\,d^*u + \log\varepsilon\right) = 0\,. \tag{2.336}$$

It will then follow that, using (2.327),

$$\int'_{\mathbb{C}} f(u)\,\frac{1}{|1-u|_{\mathbb{C}}}\,d^*u = Pfw\int f(u)\,\frac{1}{|1-u|_{\mathbb{C}}}\,d^*u\,. \tag{2.337}$$

To prove (2.336) it is enough to investigate the integral

$$j(\varepsilon) := \int_{|z|\leq 1, |1-z|\geq \varepsilon} ((1-z)(1-\bar{z}))^{-1} \, |dz \wedge d\bar{z}| \,, \tag{2.338}$$

and show that $j(\varepsilon) = \alpha \log \varepsilon + o(1)$ for $\varepsilon \to 0$. A similar statement then holds for

$$\int_{|z|\leq 1, |1-z^{-1}|\geq \varepsilon} ((1-z)(1-\bar{z}))^{-1} \, |dz \wedge d\bar{z}|. \tag{2.339}$$

Indeed, with $\delta = |1-z|$ one has $|1-z^{-1}| = \delta + O(\delta^2)$ when $\delta \to 0$ and $\log(\delta + O(\delta^2)) = \log \delta + o(1)$. Moreover the nuance between the conditions $|u| \geq \varepsilon$ and $|u|_{\mathbb{C}} \geq \varepsilon$ replaces ε by a power which only alters α in the behavior $j(\varepsilon) = \alpha \log \varepsilon + o(1)$.
One has $j(\varepsilon) = \int_D |dZ \wedge d\bar{Z}|$, where $Z = \log(1-z)$ and the domain D is the image by $z \mapsto \log(1-z)$ of $\{z \,|\, |z| \leq 1, |1-z| \geq \varepsilon\}$. It is contained in the rectangle

$$R_\varepsilon = \{Z = x + iy \,|\, \log \varepsilon \leq x \leq \log 2 \,,\, -\frac{\pi}{2} \leq y \leq \frac{\pi}{2}\} \,, \tag{2.340}$$

and is bounded by the curve $x = \log(2\cos y)$ which comes from the equation of the circle $|z| = 1$ in polar coordinates centered at $z = 1$. One thus gets

$$j(\varepsilon) = 4 \int_{\log \varepsilon}^{\log 2} \arccos(e^x/2) \, dx \,; \tag{2.341}$$

when $\varepsilon \to 0$ one has $j(\varepsilon) \sim 2\pi \log(1/\varepsilon)$, which is the area in the measure $|dZ \wedge d\bar{Z}|$ of the rectangle

$$\{Z = x + iy \,|\, \log \varepsilon \leq x \leq 0 \,,\, -\pi/2 \leq y \leq \pi/2\} \,. \tag{2.342}$$

One has $|R_\varepsilon| - 2\pi \log 2 = 2\pi \log(1/\varepsilon)$. When $\varepsilon \to 0$ the area of $R_\varepsilon \setminus D$ converges to

$$4 \int_{-\infty}^{\log 2} \arcsin(e^x/2) \, dx = -4 \int_0^{\pi/2} \log(\sin u) \, du = 2\pi \log 2 \,, \tag{2.343}$$

so that $j(\varepsilon) = 2\pi \log(1/\varepsilon) + o(1)$ when $\varepsilon \to 0$. ☐

We can now express again this finite value using the variant of the minimal subtraction scheme:

Corollary 2.42. *One has, with* $\mu = \frac{e^{-\gamma}}{2\pi}$,

$$\int_{\mathbb{C}}' f(u) \, \frac{1}{|1-u|_{\mathbb{C}}} \, d^*u = \lim_{\varepsilon \to 0} \left(\int \frac{f(u)}{|1-u|_{\mathbb{C}}^{1-\varepsilon}} \, \mu^{-2\varepsilon} \, d^*u - \frac{1}{\varepsilon} \, f(1) \right) . \tag{2.344}$$

Note the factor 2 in the exponent of μ.

8.5. Reformulation of the explicit formula.

Thus, with the above choice of basic characters for local fields, one has for any test function

$$\int_K' f(u) \, \frac{1}{|1-u|} \, d^*u = Pfw \int f(u) \, \frac{1}{|1-u|} \, d^*u \,. \tag{2.345}$$

LEMMA 2.43. *Let K be a local field, α_0 a normalized character as above and α, $\alpha(x) = \alpha_0(\lambda x)$ an arbitrary character of K. Let $\int'$ be defined as in Theorem 2.36 relative to α. Then, for any test function f,*

$$\int_K' f(u) \frac{1}{|1-u|} d^*u = \log|\lambda| f(1) + Pfw \int f(u) \frac{1}{|1-u|} d^*u \,. \tag{2.346}$$

PROOF. The lemma follows from (2.345) combined with (2.100). □

Let us now pass to the global case. Recall that if α, $\alpha \neq 1$, is a character of the additive group of adèles $\mathbb{A}_{\mathbb{K}}$ such that $\alpha|_{\mathbb{K}} = 1$, there exists a differental idèle $d = (d_v)$ such that, (cf. [**293**])

$$\alpha_v(x) = \alpha_{0,v}(d_v\, x) \tag{2.347}$$

where $\alpha = \prod \alpha_v$ and each local character $\alpha_{0,v}$ is normalized as above.
We rewrite the Weil formula (Theorem 2.37 and (2.311)) in the following form

THEOREM 2.44. *Let $\mathbb{K}$ be a global field, α a nontrivial character of $\mathbb{A}_{\mathbb{K}}/\mathbb{K}$ and $\alpha = \prod \alpha_v$ its local factors. Let $h \in \mathcal{S}(C_{\mathbb{K}})$ have compact support. Then*

$$\hat{h}(0) + \hat{h}(1) - \sum_{\chi \in \widehat{C_{\mathbb{K},1}}} \sum_{Z_{\tilde{\chi}}} \hat{h}(\tilde{\chi}, \rho) = \sum_v \int_{\mathbb{K}_v^*}' \frac{h(u^{-1})}{|1-u|} d^*u \tag{2.348}$$

where $\int'$ is normalized by α_v as in Theorem 2.36, and

$$\hat{h}(\chi, z) = \int h(u)\, \chi(u)\, |u|^z\, d^*u.$$

PROOF. This follows from (2.311), Lemma 2.43 and the equality $\log|d| = \sum_v \log|d_v|$. □

9. Spectral realization of critical zeros of L-functions

In this section we describe the spectral realization of critical zeros of L-functions given in [**71**]. Let $\mathbb{K}$ be a global field and $\mathbb{A}_{\mathbb{K}}$ the ring of adèles, $\mathbb{A}_{\mathbb{K}}^*$ the group of idèles . The *idèles class group* is the quotient $C_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$ of the group of idèles by the multiplicative group $\mathbb{K}^*$. By construction this group acts by multiplication on the following space.

DEFINITION 2.45. *The adèle class space of a global field $\mathbb{K}$ is the quotient $X_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$.*

As remarked already in the case of $\mathbb{Q}$ the action of $\mathbb{K}^*$ on $\mathbb{A}_{\mathbb{K}}$ is not proper and the adèle class space $X_{\mathbb{K}}$ is a prototypical example of a quotient space which can only be properly understood with the tools of noncommutative geometry. This will be done in Chapter 4 §5.1. In this section we show that, even using a naive approach, one can define Sobolev spaces $L^2_\delta(X_{\mathbb{K}})_0$ of functions on $X_{\mathbb{K}}$ and obtain a spectral realization of the critical zeros of L-functions associated to Grössencharakters without having to define the L-functions, let alone their analytic continuation, before stating the result. We restrict to global fields of characteristic zero i.e. to number fields to simplify the notations but the result makes sense in general. We fix as above, in §8.3, a non-canonical isomorphism

$$C_{\mathbb{K}} = C_{\mathbb{K},1} \times \mathbb{R}_+^*\,, \quad C_{\mathbb{K},1} = \mathrm{Ker}(\mathrm{Mod}) \tag{2.349}$$

where Mod is the module of the group $C_{\mathbb{K}}$. In general one replaces $\mathbb{R}_+^*$ by the range N of the module. Using (2.349) one extends the characters χ of the compact group $C_{\mathbb{K},1}$ as characters $\tilde{\chi}$ of $C_{\mathbb{K}}$, trivial on $1 \times \mathbb{R}_+^*$.

To define the Sobolev spaces $L^2_\delta(X_{\mathbb{K}})_0$ we first consider the codimension 2 subspace of the Bruhat-Schwartz space $\mathcal{S}(\mathbb{A}_{\mathbb{K}})$ (cf. [**36**]) given by the conditions

$$f(0) = 0\,, \quad \int f\,dx = 0\,. \tag{2.350}$$

On this subspace $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$ we put the inner product corresponding to the norm

$$||f||^2 = \int_{C_{\mathbb{K}}} |\mathfrak{E}(f)(x)|^2\,(1 + \log|x|^2)^{\delta/2}\,d^*x\,. \tag{2.351}$$

where

$$\mathfrak{E}(f)\,(g) = |g|^{1/2} \sum_{q \in \mathbb{K}^*} f(qg) \quad \forall\, g \in C_{\mathbb{K}}\,. \tag{2.352}$$

We let $L^2_\delta(X)_0$ be the separated completion of $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$ for the inner product given by (2.351). We let ϑ_{a} be the representation of $C_{\mathbb{K}}$ on $\mathcal{S}(\mathbb{A}_{\mathbb{K}})$ given by

$$(\vartheta_{\mathrm{a}}(\alpha)\xi)(x) = \xi(\alpha^{-1}x) \quad \forall\,\alpha \in C_{\mathbb{K}}\,,\ x \in \mathbb{A}_{\mathbb{K}}\,. \tag{2.353}$$

The Sobolev norm on $L^2_\delta(C_{\mathbb{K}})$ is defined as

$$||\xi||^2 = \int_{C_{\mathbb{K}}} |\xi(x)|^2\,(1 + \log|x|^2)^{\delta/2}\,d^*x\,. \tag{2.354}$$

The linear map $\mathfrak{E} : \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0 \to L^2_\delta(C_{\mathbb{K}})$ satisfies

$$\|\mathfrak{E}(f)\|^2_\delta = \|f\|^2_\delta \tag{2.355}$$

by construction. Thus it extends to an isometry, still denoted $\mathfrak{E}$,

$$\mathfrak{E} : L^2_\delta(X)_0 \hookrightarrow L^2_\delta(C_{\mathbb{K}})\,. \tag{2.356}$$

One has

$$\begin{aligned}
\mathfrak{E}(\vartheta_{\mathrm{a}}(\alpha)f)(g) &= |g|^{1/2} \sum_{\mathbb{K}^*} (\vartheta_{\mathrm{a}}(\alpha)f)(qg) = |g|^{1/2} \sum_{\mathbb{K}^*} f(\alpha^{-1}qg) \\
&= |g|^{1/2} \sum_{\mathbb{K}^*} f(q\alpha^{-1}g) = |\alpha|^{1/2}\,|\alpha^{-1}g|^{1/2} \sum_{\mathbb{K}^*} f(q\alpha^{-1}g) \\
&= |\alpha|^{1/2}(\vartheta_{\mathrm{m}}(\alpha)\,\mathfrak{E}(f))(g)
\end{aligned}$$

so that

$$\mathfrak{E}\,\vartheta_{\mathrm{a}}(\alpha) = |\alpha|^{1/2}\,\vartheta_{\mathrm{m}}(\alpha)\,\mathfrak{E}\,. \tag{2.357}$$

The equality (2.357) shows that the natural representation ϑ_{a} of $C_{\mathbb{K}}$ on $L^2_\delta(X)_0$ corresponds by the isometry $\mathfrak{E}$ to the restriction of $|\alpha|^{1/2}\,\vartheta_{\mathrm{m}}(\alpha)$ to the invariant subspace given by the range of $\mathfrak{E}$.

DEFINITION 2.46. *We let $\mathcal{H}$ be the cokernel of the map $\mathfrak{E}$ and $\underline{\vartheta}_{\mathrm{m}}$ the quotient representation of $C_{\mathbb{K}}$ on $\mathcal{H}$. For χ a character of $C_{\mathbb{K},1}$ we let*

$$\mathcal{H}_\chi = \{\xi \in \mathcal{H} \mid \underline{\vartheta}_{\mathrm{m}}(\gamma)\,\xi = \chi\,(\gamma)\,\xi \ , \ \forall\,\gamma \in C_{\mathbb{K},1}\} \tag{2.358}$$

Since $C_{\mathbb{K},1}$ is a compact group it follows that the Hilbert space $\mathcal{H}$ splits as a direct sum

$$\mathcal{H} = \bigoplus_{\chi \in \hat{C}_{\mathbb{K},1}} \mathcal{H}_\chi \tag{2.359}$$

and the representation $\underline{\vartheta}_{\mathrm{m}}$ also decomposes as a direct sum of representations $\underline{\vartheta}_{\mathrm{m}\chi}$. We then restrict $\underline{\vartheta}_{\mathrm{m}\chi}$ to the subgroup $1 \times \mathbb{R}_+^* \subset C_{\mathbb{K}}$ and let D_χ be the infinitesimal generator of this restriction, i.e.

$$D_\chi\,\xi = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\,(\underline{\vartheta}_{\mathrm{m}\chi}(e^\varepsilon) - 1)\,\xi \tag{2.360}$$

The spectral realization of critical zeros of L-functions is then given by:

THEOREM 2.47. *Let χ, $\delta > 1$, $\mathcal{H}_\chi$ and D_χ be as above. Then D_χ has discrete spectrum, $\mathrm{Sp}\,D_\chi \subset i\,\mathbb{R}$ is the set of imaginary parts of zeros of the L-function with Grössencharakter $\tilde{\chi}$ which have real part equal to $\frac{1}{2}$; $\rho \in \mathrm{Sp}\,D \Leftrightarrow L\left(\tilde{\chi}, \frac{1}{2} + \rho\right) = 0$ and $\rho \in i\,\mathbb{R}$. Moreover the multiplicity of ρ in $\mathrm{Sp}\,D$ is equal to the largest integer $n < 1 + \frac{\delta}{2}$, $n \leq$ multiplicity of $\frac{1}{2} + \rho$ as a zero of L.*

Theorem 2.47 has a similar formulation when the characteristic of $\mathbb{K}$ is nonzero. The following corollary is valid for global fields $\mathbb{K}$ of arbitrary characteristic, with $N = \mathrm{Mod}(C_{\mathbb{K}})$ in general.

COROLLARY 2.48. *For any Schwartz function $h \in \mathcal{S}(C_{\mathbb{K}})$ the operator $\underline{\vartheta}_{\mathrm{m}}(h) = \int \underline{\vartheta}_{\mathrm{m}}(g)\,h(g)\,d^*\,g$ in $\mathcal{H}$ is of trace class, and its trace is given by*

$$\mathrm{Tr}\,\underline{\vartheta}_{\mathrm{m}}(h) = \sum_{\substack{L\left(\tilde{\chi}, \frac{1}{2}+\rho\right)=0 \\ \rho \in i\,\mathbb{R}/N^\perp}} \hat{h}(\tilde{\chi}, \rho)$$

where the multiplicity is counted as in Theorem 2.47 and where the Fourier transform $\hat{h}$ of h is defined by

$$\hat{h}(\tilde{\chi}, \rho) = \int_{C_{\mathbb{K}}} h(u)\,\tilde{\chi}(u)\,|u|^\rho\,d^*\,u\,.$$

9.1. L-functions and homogeneous distributions on $\mathbb{A}_{\mathbb{K}}$.

Let us recall from [**296**] how the L-function $L(\chi_0, s)$ appears as a normalization factor for homogeneous distributions on $\mathbb{A}_{\mathbb{K}}$.

First let K be a local field and χ a quasi-character of K^*,

$$\chi(x) = \chi_0(x)\,|x|^s \quad , \quad \chi_0 : K^* \to U(1)\,. \tag{2.361}$$

DEFINITION 2.49. *A distribution D on K is homogeneous of weight χ if one has*

$$\langle f^a, D\rangle = \chi(a)^{-1}\,\langle f, D\rangle \tag{2.362}$$

for all test functions f and all a in K^*, where by definition

$$f^a(x) = f(ax)\,. \tag{2.363}$$

When $\sigma = \Re(s) > 0$, there exists up to normalization only one homogeneous distribution of weight χ on K (cf. [**296**]). It is given by the absolutely convergent integral

$$\int_{K^*} f(x)\,\chi(x) d^*x = \Delta_\chi(f)\,. \tag{2.364}$$

In particular, let K be non-Archimedean. Then, for any compactly supported locally constant function f on K, one has

$$f(x) - f(\pi^{-1}x) = 0 \quad \forall x,\ |x| \le \delta\,. \tag{2.365}$$

Thus, for any $s \in \mathbb{C}$ the integral

$$\int_{K^*} (f(x) - f(\pi^{-1}x))\,|x|^s\, d^*x = \Delta'_s(f) \tag{2.366}$$

with the multiplicative Haar measure d^*x normalized by

$$\langle \mathbf{1}_{\mathcal{O}^*}, d^*x\rangle = 1\,. \tag{2.367}$$

defines a distribution on K with the properties

$$\langle \mathbf{1}_{\mathcal{O}}, \Delta'_s\rangle = 1 \tag{2.368}$$

$$\langle f^a, \Delta'_s\rangle = |a|^{-s}\,\langle f, \Delta'_s\rangle \tag{2.369}$$

and

$$\Delta'_s = (1 - q^{-s})\,\Delta_s, \tag{2.370}$$

where $|\pi| = q^{-1}$. Let us check (2.368)–(2.370). With $f = \mathbf{1}_{\mathcal{O}}$ one has $f(\pi^{-1}x) = 1$ iff $\pi^{-1}x \in \mathcal{O}$, i.e., $x \in \pi\mathcal{O} = \mathfrak{m}$. Thus $\Delta'_s(\mathbf{1}_{\mathcal{O}}) = \int_{\mathcal{O}^*} d^*x = 1$. Let us check (2.370). One has $\int f(\pi^{-1}x)\,|x|^s\, d^*x = \int f(y)\,|\pi|^s\,|y|^s\, d^*y = |\pi|^s\,\Delta_s(f)$. But $|\pi| < 1$, $|\pi| = \frac{1}{q}$. Note also that for $s = 1$ and $f = \mathbf{1}_{\mathcal{O}}$, one gets $\int_{\mathcal{O}^*} dx = \left(1 - \frac{1}{q}\right)\int_{\mathcal{O}} dx$.

Let then χ be a quasi-character of $C_{\mathbb{K}}$ and write as above

$$\chi = \prod \chi_v\,, \qquad \chi(x) = \chi_0(x)\,|x|^s \tag{2.371}$$

where $s \in \mathbb{C}$ and χ_0 is a character. Let P be the finite set of finite places where it is ramified. For any finite place $v \notin P$, let $\Delta'_v(s)$ be the unique homogeneous distribution of weight χ_v normalized by

$$\langle \Delta'_v(s), \mathbf{1}_{\mathcal{O}_v}\rangle = 1\,. \tag{2.372}$$

For any $v \in P$ or any infinite place and for $\sigma = \Re(s) > 0$ let Δ'_v be given by (2.364) which is homogeneous of weight χ_v but unnormalized. Then the infinite tensor product

$$\Delta'_s = \prod \Delta'_v(s) \tag{2.373}$$

makes sense as a continuous linear form on $\mathcal{S}(\mathbb{A}_{\mathbb{K}})$ and is homogeneous of weight χ. This homogeneous distribution is not equal to 0 since $\Delta'_v \neq 0$ for any $v \in P$ and any infinite place as well. It is finite by construction of the space $\mathcal{S}(\mathbb{A}_{\mathbb{K}})$ of test functions as an infinite tensor product,

$$\mathcal{S}(\mathbb{A}_{\mathbb{K}}) = \otimes\,(\mathcal{S}(\mathbb{K}_v), \mathbf{1}_{\mathcal{O}_v})\,. \tag{2.374}$$

LEMMA 2.50. *(cf. [**296**]) For $\sigma = \Re(s) > 1$, the following integral converges absolutely:*

$$\int f(x)\,\chi_0(x)\,|x|^s\,d^*x = \Delta_s(f) \quad \forall\, f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}}) \tag{2.375}$$

and $\Delta_s(f) = L(\chi_0, s)\,\Delta'_s(f)$.

PROOF. To get the absolute convergence, one can assume that $f = \mathbf{1}_{\mathcal{O}}$ and $\chi_0 = 1$. Then one has to control an infinite product of local terms, given locally for the Haar measure d^*x on $\mathbb{K}_v^*$ such that $\int_{\mathcal{O}_v^*} d^*x = 1$, by

$$\int_{\mathcal{O}\cap\mathbb{K}_v^*} |x|^s\,d^*x \quad (s \text{ real}) \tag{2.376}$$

which is $1 + q_v^{-s} + q_v^{-2s} + \cdots = (1 - q_v^{-s})^{-1}$. Thus the convergence for $\sigma = \Re(s) > 1$ is the same as for the zeta function.

To prove the second equality, one only needs to consider the infinite tensor product for finite places $v \notin P$. Then by (2.370) one has $\Delta'_v = (1 - q_v^{-\alpha_v})\,\Delta_v$ where

$$q_v^{-\alpha_v} = \chi_v(\pi) = \chi_{0,v}(\pi)\,q_v^{-s} \tag{2.377}$$

with $|\pi| = q_v^{-1}$.

Thus one gets $\Delta_s = \left(\prod_{\substack{v \text{ finite} \\ v\notin P}} (1 - \chi_{0,v}(\pi)\,q_v^{-s})^{-1}\right)\,\Delta'_s = L(\chi_0, s)\,\Delta'_s.$

By construction Δ'_s makes sense whenever $\sigma > 0$ and is a holomorphic function of s (for fixed f). Let us review briefly (cf. [**296**]) how to extend the definition of Δ_s.

We let as above $\mathbb{K}$ be a global field, we fix a nontrivial additive character α of $\mathbb{A}_{\mathbb{K}}$, trivial on $\mathbb{K}$,

$$\alpha(x + y) = \alpha(x)\,\alpha(y) \in U(1)\ ,\ \alpha(q) = 1 \quad \forall\, q \in \mathbb{K}\,. \tag{2.378}$$

We then identify the dual of the locally compact additive group $\mathbb{A}_{\mathbb{K}}$ with $\mathbb{A}_{\mathbb{K}}$ itself by the pairing

$$\langle x, y\rangle = \alpha(xy)\,. \tag{2.379}$$

One shows (cf. [**293**]) that the lattice $\mathbb{K} \subset \mathbb{A}_{\mathbb{K}}$, i.e., the discrete and cocompact additive subgroup $\mathbb{K}$, is its own dual

$$\langle x, q\rangle = 1 \quad \forall\, q \in \mathbb{K} \quad \Leftrightarrow \quad x \in \mathbb{K}\,. \tag{2.380}$$

Since $\mathbb{A}_{\mathbb{K}}$ is the restricted product of the local fields $\mathbb{K}_v$, one can write α as an infinite product

$$\alpha = \prod \alpha_v \tag{2.381}$$

where for almost all v, one has $\alpha_v = 1$ on $\mathcal{O}_v$. Let us recall the definition of the space $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$. We have

$$\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0 = \{f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}}) \mid f(0) = 0\ ,\ \int f\,dx = 0\}\,. \tag{2.382}$$

LEMMA 2.51. *Let $f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$. Then the series*

$$\mathfrak{E}(f)\,(g) = |g|^{1/2} \sum_{q \in \mathbb{K}^*} f(qg) \quad \forall\, g \in C_{\mathbb{K}} \tag{2.383}$$

converges absolutely and one has

$$\forall\, n\,,\ \exists\, c, \quad |\mathfrak{E}(f)(g)| \leq c\, e^{-n|\log|g||} \quad \forall\, g \in C_{\mathbb{K}} \tag{2.384}$$

and $\mathfrak{E}(\widehat{f})\,(g) = \mathfrak{E}(f)\,(g^{-1})$.
For $\mathbb{K}$ of characteristic zero and for a suitable decomposition $C_{\mathbb{K}} = C_{\mathbb{K},1} \times \mathbb{R}_+^*$ the range of $\mathfrak{E}$ is stable under the generator $\lambda \frac{\partial}{\partial \lambda}$ of the scaling group.

PROOF. Let us first recall the formal definition ([**36**]) of the Bruhat-Schwartz space $\mathcal{S}(G)$ for an arbitrary locally compact abelian group G. One considers all pairs of subgroups G_1, G_2 of G such that G_1 is generated by a compact neighborhood of 0 in G, while G_2 is a compact subgroup of G_1 such that the quotient group is elementary, i.e., is of the form $\mathbb{R}^a\, \mathbb{T}^b\, \mathbb{R}^c\, F$ for F a finite group. By definition the Bruhat-Schwartz space $\mathcal{S}(G)$ is the inductive limit of the Schwartz spaces $\mathcal{S}(G_1/G_2)$, where the latter have the usual definition in terms of rapid decay of all derivatives. Since G_1 is open in G, any element of $\mathcal{S}(G_1/G_2)$ extended by 0 outside G_1 defines a continuous function on G. By construction $\mathcal{S}(G)$ is the union of the subspaces $\mathcal{S}(G_1/G_2)$ and it is endowed with the inductive limit topology.

Let $\widehat{G}$ be the Pontrjagin dual of G. Then the Fourier transform, which depends upon the normalization of the Haar measure on G, gives an isomorphism of $\mathcal{S}(G)$ with $\mathcal{S}(\widehat{G})$.

Let Γ be a lattice in the locally compact abelian group G. Then any function $f \in \mathcal{S}(G)$ is admissible for the pair G, Γ in the sense of [**293**], and the Poisson summation formula (cf. [**293**]) is the equality

$$\mathrm{Covol}\,(\Gamma) \sum_{\gamma \in \Gamma} f(\gamma) = \sum_{\beta \in \Gamma^{\perp}} \widehat{f}(\beta)\,, \tag{2.385}$$

where $\Gamma^{\perp}$ is the dual of the lattice Γ, and

$$\widehat{f}(\beta) = \int f(a)\, \beta(a)\, da\,. \tag{2.386}$$

Both sides of (2.385) depend upon the normalization of the Haar measure on G.

In our case we let $\mathbb{A}_{\mathbb{K}}$ be as above the additive group of adèles on $\mathbb{K}$. We normalize the additive Haar measure dx on $\mathbb{A}_{\mathbb{K}}$ by

$$\mathrm{Covol}(\mathbb{K}) = 1\,. \tag{2.387}$$

We then take $\Gamma = xk$, for some $x \in \mathbb{A}_{\mathbb{K}}^{-1}$. One has

$$\mathrm{Covol}\,(xk) = |x|\,. \tag{2.388}$$

The dual $\Gamma^{\perp}$ of the lattice xk, for x invertible in $\mathbb{A}_{\mathbb{K}}$, is the lattice $\Gamma^{\perp} = x^{-1}\mathbb{K}$. Thus the Poisson formula (2.385) reads, for any $f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})$,

$$|x| \sum_{q \in \mathbb{K}} f(xq) = \sum_{q \in \mathbb{K}} \widehat{f}(x^{-1}q)\,. \tag{2.389}$$

This we can rewrite as

$$|x| \sum_{\mathbb{K}^*} f(xq) = \sum_{\mathbb{K}^*} \widehat{f}(x^{-1}q) + \delta \tag{2.390}$$

with $\delta = -|x|\, f(0) + \int f(y)\, dy$.

We can then rewrite (2.390) as the equality

$$\mathfrak{E}(f)(x) = \mathfrak{E}(\widehat{f})\left(\frac{1}{x}\right), \tag{2.391}$$

valid for all $f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$. It remains to control the growth of $\mathfrak{E}(f)(x)$ on $C_{\mathbb{K}}$, but by (2.391), it is enough to understand what happens for $|x|$ large.

We only treat the case of number fields; the general case is similar. Let $\mathbb{A}_{\mathbb{K}} = \mathbb{A}_{\mathbb{K},f} \times \mathbb{A}_{\mathbb{K},\infty}$ be the decomposition of the ring of adèles corresponding to finite and Archimedean places. Thus $\mathbb{A}_{\mathbb{K},\infty} = \prod_{\Sigma_\infty} \mathbb{K}_v$ where Σ_∞ is the set of Archimedean places.

Any element of $\mathcal{S}(\mathbb{A}_{\mathbb{K}})$ is a finite linear combination of test functions

$$f = f_0 \otimes f_1 \tag{2.392}$$

where $f_0 \in \mathcal{S}(\mathbb{A}_{\mathbb{K},f})$, $f_1 \in \mathcal{S}(\mathbb{A}_{\mathbb{K},\infty})$ (cf. [**295**] 39); thus it is enough to control the growth of $\mathfrak{E}(f)(x)$ for such f and $|x|$ large.

Let $\mathbb{A}^*_{\mathbb{K},1} = \{x \in \mathbb{A}^*_{\mathbb{K}}; |x| = 1\}$ be the group of idèles of module one. Since $\mathbb{A}^*_{\mathbb{K},1}/\mathbb{K}^*$ is compact (cf. [**293**]), we shall fix a compact subset K_1 of $\mathbb{A}^*_{\mathbb{K},1}$ whose image in $\mathbb{A}^*_{\mathbb{K},1}/\mathbb{K}^*$ is this compact group.

Let $\mathbf{d}$ be the diagonal embedding

$$\mathbf{d} \,:\, \lambda \in \mathbb{R}^*_+ \mapsto (\lambda, \ldots, \lambda) \in \prod_{\Sigma_\infty} \mathbb{K}^*_v\,. \tag{2.393}$$

This yields an isomorphism

$$\mathbb{A}^*_{\mathbb{K}} = \mathbb{A}^*_{\mathbb{K},1} \times \operatorname{Range} \mathbf{d}\,. \tag{2.394}$$

One has $f_0 \in \mathcal{S}(\mathbb{A}_{\mathbb{K},f})$, hence (cf. [**295**]), $f_0 \in C_c(\mathbb{A}_{\mathbb{K},f})$ and we let $K_0 = \text{Support}\, f_0$. Since K_0 is compact, one can find a finite subset P of the set of finite places, and $C < \infty$ such that

$$y \in K = (K_f)^{-1} K_0 \Rightarrow |y_v| \leq 1 \quad \forall\, v \notin P, \quad |y_v| \leq C \quad \forall\, v\,, \tag{2.395}$$

where K_f is the projection of K_1 on $\mathbb{A}_{\mathbb{K},f}$.

We let Ω be the compact open subgroup of $\mathbb{A}_{\mathbb{K},f}$ determined by

$$|a_v| \leq 1 \quad \forall\, v \notin P, \quad |a_v| \leq C \quad \forall\, v\,. \tag{2.396}$$

By construction $\mathfrak{E}(f)(x)$ only depends upon the class of x in $\mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*$. Thus, to control the behaviour of $\mathfrak{E}(f)(x)$ for $|x| \to \infty$, we can take $x = (x_f, x_\infty) \in K_1$ and consider $\mathfrak{E}(f)(\lambda\, x)$ for $\lambda \in \mathbb{R}^*_+\,, \lambda \to \infty$. Now let $q = (q_f, q_\infty) \in \mathbb{K}$. Then, we have

$$f(q\, \lambda\, x) = f_0(q_f\, x_f)\, f_1(q_\infty\, \lambda\, x_\infty) \tag{2.397}$$

and this vanishes unless $q_f\, x_f \in K_0$, i.e., unless $q_f \in K$. Then by (2.395) one has $q_f \in \Omega$. Let Γ be the lattice in $\prod_{\Sigma_\infty} \mathbb{K}_v$ determined by

$$\Gamma = \{q_\infty \,|\, q \in \mathbb{K},\, q_f \in \Omega, \}\,. \tag{2.398}$$

The size of $\mathfrak{E}(f)(\lambda\, x)$ is hence controlled (up to the square root of $|\lambda\, x|$) by

$$C \sum_{n \in \Gamma^*} |f_1(\lambda\, x_\infty\, n)|\,, \tag{2.399}$$

where x_∞ varies in the projection K_∞ of K_1 on $\prod_{\Sigma_\infty} \mathbb{K}_v^*$.

Since $f_1 \in \mathcal{S}(\mathbb{A}_{\mathbb{K},\infty})$, this shows that $\mathfrak{E}(f)(x)$ decays faster than any power of $|x|$ for $|x| \to \infty$.

We have shown that $\mathfrak{E}(f)$ has rapid decay in terms of $|x|$, for $|x| \to \infty$. Using (2.391) and the stability of $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$ under Fourier transform, we see that it also has exponential decay in terms of $|\log|x||$ when $|\log|x|| \to \infty$.

The second statement follows from the stability of $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$ under the generator of the scaling group (2.393). $\square$

We then get

LEMMA 2.52. *(cf. [**296**]) For $\sigma = \Re(s) > 0$, and any character χ_0 of $C_{\mathbb{K}}$, one has*

$$\int \mathfrak{E}(f)(x)\, \chi_0(x)\, |x|^{s-1/2}\, d^*x = cL(\chi_0, s)\, \Delta'_s(f) \quad \forall\, f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0 \tag{2.400}$$

where the nonzero constant c depends upon the normalization of the Haar measure d^*x on $C_{\mathbb{K}}$.

PROOF. For $\sigma = \Re(s) > 1$, the equality follows from Lemma 2.50, but since both sides are analytic in s it holds in general. $\square$

As in Lemma 2.50, we continue to use the notation $\Delta_s(f)$ for $\sigma = \Re(s) > 0$.

9.2. Approximate units in the Sobolev spaces $L^2_\delta(C_{\mathbb{K}})$.

We first consider, for $\delta > 1$, the Hilbert space $L^2_\delta(\mathbb{R})$ of functions $\xi(u)$, $u \in \mathbb{R}$, with squared norm given by

$$\int_{\mathbb{R}} |\xi(u)|^2\, (1+u^2)^{\delta/2}\, du\,. \tag{2.401}$$

We let $\rho(u) = (1+u^2)^{\delta/2}$. It is comparable to $(1+|u|)^\delta$ and in particular,

$$\frac{\rho(u+a)}{\rho(u)} \le c\, \rho(a) \quad \forall\, u \in \mathbb{R}\,,\; a \in \mathbb{R} \tag{2.402}$$

with $c = 2^{\delta/2}$. We then let $\lambda(v)$ be the translation operator

$$(\lambda(v)\xi)(u) = \xi(u-v) \quad \forall\, u, v \in \mathbb{R}\,. \tag{2.403}$$

One has $\int_{\mathbb{R}} |\xi(u-v)|^2\, \rho(u)\, du = \int_{\mathbb{R}} |\xi(u)|^2\, \rho(u+v)\, du$ so that by (2.402) it is less than $c \int_{\mathbb{R}} |\xi(u)|^2\, \rho(u)\, \rho(v)\, du = c\, \rho(v)\, \|\xi\|^2$,

$$\|\lambda(v)\| \le (c\, \rho(v))^{1/2}\,. \tag{2.404}$$

This shows that $\lambda(f) = \int f(v)\,\lambda(v)\,dv$ makes sense as soon as

$$\int |f(v)|\,\rho(v)^{1/2}\,dv < \infty\,. \tag{2.405}$$

This holds for all $f \in \mathcal{S}(\mathbb{R})$.

LEMMA 2.53. *There exists an approximate unit $f_n \in \mathcal{S}(\mathbb{R})$, such that the Fourier transform of f_n has compact support, $\|\lambda(f_n)\| \leq C \ \ \forall\, n$, and*

$$\lambda(f_n) \to 1 \textit{ strongly in } L^2_\delta(\mathbb{R})\,. \tag{2.406}$$

PROOF. Let f be a function, $f \in \mathcal{S}(\mathbb{R})$, whose Fourier transform $\widehat{f}(s) = \int f(x)e^{ixs}dx$ has compact support, and such that $\int f\,dx = 1$ (i.e. $\widehat{f}(0) = 1$). Let then

$$f_n(v) = n\,f(nv)\,, \quad n = 1, 2, \ldots \tag{2.407}$$

One has $\int |f_n(v)|\,\rho(v)^{1/2}\,dv = \int |f(u)|\,\rho\left(\frac{u}{n}\right)^{1/2}\,du \leq \int |f(u)|\,\rho(u)^{1/2}\,du$. Thus $\|\lambda(f_n)\|$ is uniformly bounded. We can assume that $\widehat{f}$ is equal to 1 on $[-1,1]$. Then $\widehat{f_n}$ is equal to 1 on $[-n,n]$ and $\lambda(f_n)\xi = \xi$ for any ξ with $\operatorname{Supp}\widehat{\xi} \subset [-n,n]$. By uniformity one gets that $\lambda(f_n) \to 1$ strongly. ☐

Let us now identify the dual $(L^2_\delta)^*$ of the Hilbert space L^2_δ with $L^2_{-\delta}$ by means of the pairing

$$\langle \xi, \eta \rangle_0 = \int_{\mathbb{R}} \xi(u)\,\eta(u)\,du\,. \tag{2.408}$$

Since L^2_δ is a Hilbert space, it is its own dual using the pairing

$$\langle \xi, \eta_1 \rangle = \int_{\mathbb{R}} \xi(u)\,\eta_1(u)(1+u^2)^{\delta/2}\,du\,. \tag{2.409}$$

If we let $\eta(u) = \eta_1(u)(1+u^2)^{\delta/2}$, then

$$\int |\eta_1(u)|^2\,(1+u^2)^{\delta/2}\,du = \int |\eta(u)|^2\,(1+u^2)^{-\delta/2}\,du \tag{2.410}$$

which is the natural norm square for $L^2_{-\delta}$.

Given a quasicompact group such as $C_{\mathbb{K}}$ with module

$$|\ .\ | : C_{\mathbb{K}} \to \mathbb{R}^*_+\,, \tag{2.411}$$

we let d^*g be the Haar measure on $C_{\mathbb{K}}$ normalized by

$$\int_{|g|\in[1,\Lambda]} d^*g \sim \log\Lambda\,, \quad \Lambda \to \infty\,, \tag{2.412}$$

and we let $L^2_\delta(C_{\mathbb{K}})$ be defined by the norm

$$\int_{C_{\mathbb{K}}} |\xi(g)|^2\,(1+\log|g|^2)^{\delta/2}\,d^*g\,. \tag{2.413}$$

It is, when the module of $\mathbb{K}$ is $\mathbb{R}^*_+$, a direct sum of spaces (2.401), labeled by the characters χ_0 of the compact group

$$C_{\mathbb{K},1} = \operatorname{Ker}\operatorname{Mod}\,. \tag{2.414}$$

The pairing between $L^2_\delta(C_{\mathbb{K}})$ and $L^2_{-\delta}(C_{\mathbb{K}})$ is given by

$$\langle \xi, \eta\rangle = \int \xi(g)\, \eta(g)\, d^* g\,. \tag{2.415}$$

The regular representation $\vartheta_{\rm m}$ of $C_{\mathbb{K}}$ in $L^2_\delta(C_{\mathbb{K}})$ is given by

$$(\vartheta_{\rm m}(a)\xi)(g) = \xi(a^{-1}g) \quad \forall\, g, a \in C_{\mathbb{K}}\,. \tag{2.416}$$

It is not unitary but by (2.404) one has

$$\|\vartheta_{\rm m}(g)\| = O\,|\log|g||^{\delta/2}\,,\ \ |\log|g|| \to \infty\,. \tag{2.417}$$

Finally, one has, using Lemma 2.53 and the decomposition $C_{\mathbb{K}} = C_{\mathbb{K},1}\times N$.

Lemma 2.54. *There exists an approximate unit $f_n \in \mathcal{S}(C_{\mathbb{K}})$, such that $\widehat{f_n}$ has compact support, $\|\vartheta_{\rm m}(f_n)\| \le C \ \ \forall\, n$, and*

$$\vartheta_{\rm m}(f_n) \to 1 \text{ strongly in } L^2_\delta(C_{\mathbb{K}})\,. \tag{2.418}$$

9.3. Proof of Theorem 2.47.

In order to understand Range $\mathfrak{E}$ we consider its orthogonal in the dual space $L^2_{-\delta}(C_{\mathbb{K}})$. The compact subgroup

$$C_{\mathbb{K},1} = \{g \in C_{\mathbb{K}}\,;\ |g| = 1\} \tag{2.419}$$

acts by the representation $\vartheta_{\rm m}$ which is unitary when restricted to $C_{\mathbb{K},1}$. Thus one can decompose $L^2_\delta(C_{\mathbb{K}})$ and its dual $L^2_{-\delta}(C_{\mathbb{K}})$ into the direct sum of the subspaces

$$L^2_{\delta,\chi_0} = \{\xi \in L^2_\delta(C_{\mathbb{K}})\,;\ \xi(\gamma^{-1}g) = \chi_0(\gamma)\,\xi(g) \ \ \forall\, g \in C_{\mathbb{K}}\,,\ \gamma \in C_{\mathbb{K},1}\} \tag{2.420}$$

and

$$L^2_{-\delta,\chi_0} = \{\xi \in L^2_{-\delta}(C_{\mathbb{K}})\,;\ \xi(\gamma\, g) = \chi_0(\gamma)\,\xi(g) \ \ \forall\, g \in C_{\mathbb{K}}\,,\ \gamma \in C_{\mathbb{K},1}\}\,. \tag{2.421}$$

These correspond to the projections $P_{\chi_0} = \int \overline{\chi_0}(\gamma)\,\vartheta_{\rm m}(\gamma)\,d_1\gamma$ and $P^t_{\chi_0} = \int \overline{\chi_0}(\gamma)\,\vartheta_{\rm m}(\gamma)^t\,d_1\gamma$ for L^2_δ and for the dual space $L^2_{-\delta}$. In (2.421) we used the formula

$$(\vartheta_{\rm m}(\gamma)^t\,\eta)(x) = \eta(\gamma x) \tag{2.422}$$

which follows from the definition of the transpose, $\langle\vartheta_{\rm m}(\gamma)\xi,\eta\rangle = \langle\xi,\vartheta_{\rm m}(\gamma)^t\eta\rangle$ using

$$\int \xi(\gamma^{-1}x)\,\eta(x)\,d^*x = \int \xi(y)\,\eta(\gamma y)\,d^*y\,. \tag{2.423}$$

In these formulas one only uses the character χ_0 as a character of the compact subgroup $C_{\mathbb{K},1}$ of $C_{\mathbb{K}}$. One now chooses, noncanonically, an extension $\widetilde{\chi}_0$ of χ_0 as a character of $C_{\mathbb{K}}$

$$\widetilde{\chi}_0(\gamma) = \chi_0(\gamma) \quad \forall \gamma \in C_{\mathbb{K},1}\,. \tag{2.424}$$

This choice is not unique but any two such extensions differ by a character which is principal, i.e., of the form: $\gamma \mapsto |\gamma|^{is_0}$, $s_0 \in \mathbb{R}$. We fix a factorization $C_{\mathbb{K}} = C_{\mathbb{K},1}\times\mathbb{R}^*_+$, and fix $\widetilde{\chi}_0$ as being equal to 1 on $\mathbb{R}^*_+$.

We then write any element η of $L^2_{-\delta,\chi_0}(C_{\mathbb{K}})$ in the form

$$g \in C_{\mathbb{K}} \mapsto \eta(g) = \widetilde{\chi}_0(g)\,\psi(|g|) \tag{2.425}$$

where

$$\int |\psi(|g|)^2 \, (1+(\log|g|)^2)^{-\delta/2} \, d^*g < \infty \,. \tag{2.426}$$

This vector is in the orthogonal of Range $\mathfrak{E}$ iff

$$\int \mathfrak{E}(f)(x) \, \widetilde{\chi}_0(x) \, \psi(|x|) \, d^*x = 0 \quad \forall \, f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0 \,. \tag{2.427}$$

We first proceed formally and write $\psi(|x|) = \int \widehat{\psi}(t) \, |x|^{it} \, dt$ so that the left hand side of (2.427) becomes

$$\int \int \mathfrak{E}(f)(x) \, \widetilde{\chi}_0(x) \, |x|^{it} \, \widehat{\psi}(t) \, d^*x \, dt = \int \Delta_{1/2+it}(f) \, \widehat{\psi}(t) \, dt \tag{2.428}$$

(using the notation of Lemmas 2.50 and 2.52).

Let us justify this formal manipulation; since we deal with the orthogonal of an invariant subspace, we can assume that

$$\vartheta^t_{\mathrm{m}}(h) \, \eta = \eta, \tag{2.429}$$

for some h such that $\widehat{h}$ has compact support. Indeed we can use Lemma 2.54 to only consider vectors which belong to the range of

$$\vartheta^t_{\mathrm{m}}(h) = \int h(\gamma) \, \vartheta_{\mathrm{m}}(\gamma)^t \, d^*\gamma \,, \; \widehat{h} \text{ with compact support.} \tag{2.430}$$

Then, using (2.429), the Fourier transform of the tempered distribution ψ on $\mathbb{R}^*_+$ has compact support in $\mathbb{R}$. Thus, since $\mathfrak{E}(f)(x)$ has rapid decay, the equality between (2.427) and (2.428) follows from the definition of the Fourier transform of the tempered distribution ψ on $\mathbb{R}^*_+$.

Let us now describe suitable test functions $f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$ in order to test the distribution

$$\int \Delta_{\frac{1}{2}+it} \, \widehat{\psi}(t) \, dt \,. \tag{2.431}$$

We treat the case of characteristic zero. The general case is similar. For the finite places we take

$$f_0 = \otimes_{v \notin P} \mathbf{1}_{\mathcal{O}_v} \otimes f_{\chi 0} \tag{2.432}$$

where f_{χ_0} is the tensor product over ramified places of the functions equal to 0 outside $\mathcal{O}_v^*$ and to $\overline{\chi}_{0,v}$ on $\mathcal{O}_v^*$. It follows then by the definition of Δ'_s that

$$\langle \Delta'_s, f_0 \otimes f \rangle = \int f(x) \, \chi_{0,\infty}(x) \, |x|^s \, d^*x \tag{2.433}$$

for any $f \in \mathcal{S}(\mathbb{A}_{\mathbb{K},\infty})$. Moreover if the set P of finite ramified places is not empty, one has

$$f_0(0) = 0 \,, \; \int_{\mathbb{A}_{\mathbb{K},f}} f_0(x) \, dx = 0 \tag{2.434}$$

so that $f_0 \otimes f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0 \;\; \forall \, f \in \mathcal{S}(\mathbb{A}_{\mathbb{K},\infty})$.

Now let ℓ be the number of Archimedean places of $\mathbb{K}$ and consider the map $\rho : (\mathbb{R}^*_+)^\ell \to \mathbb{R}^*_+$ given by

$$\rho(\lambda_1, \ldots, \lambda_\ell) = \lambda_1 \cdots \lambda_\ell \,. \tag{2.435}$$

As soon as $\ell > 1$ this map is not proper. Given a smooth function with compact support, $b \in C_c^\infty(\mathbb{R}_+^*)$, we need to find $a \in C_c^\infty((\mathbb{R}_+^*)^\ell)$ such that the direct image of the measure $a(x)\, d^*x$ is $b(y)\, d^*y$, where $d^*x = \prod d^*x_i$ is the product of the multiplicative Haar measures.

Equivalently, one is dealing with a finite dimensional vector space E and a linear form $L : E \to \mathbb{R}$. One is given $b \in C_c^\infty(\mathbb{R})$ and asked to lift it. One can write $E = \mathbb{R} \times E_1$ and the lift can be taken as $a = b \otimes b_1$ where $b_1 \in C_c^\infty(E_1)$, $\int b_1\, dx = 1$. Thus we can in (2.433) take a function f of the form

$$f(x) = g(x)\, \overline{\chi}_{0,\infty}\,(x)\,, \tag{2.436}$$

where the function $g \in C_c^\infty(\mathbb{A}_{\mathbb{K},\infty})$ only depends upon $(|x|_v)$, $v \in \Sigma_\infty$, and is smooth with compact support, disjoint from the closed set

$$\left\{ x \in \prod_{v \in \Sigma_\infty} \mathbb{K}_v \,;\, \exists\, v \,,\, x_v = 0 \right\}. \tag{2.437}$$

Thus, to any function $b \in C_c^\infty(\mathbb{R}_+^*)$ we can assign a test function $f = f_b$ such that for any s ($\Re s > 0$)

$$\langle \Delta'_s, f_0 \otimes f_b \rangle = \int_{\mathbb{R}_+^*} b(x)\, |x|^s\, d^*x\,. \tag{2.438}$$

By Lemma 2.52, we get

$$\begin{aligned} \left\langle \int \Delta_{\frac{1}{2}+it}\, \widehat{\psi}(t)\, dt\,,\, f_0 \otimes f_b \right\rangle &= \left\langle \int L\,(\chi_0, \frac{1}{2} + it)\, \Delta'_{\frac{1}{2}+it}\, \widehat{\psi}(t)\, dt\,,\, f_0 \otimes f_b \right\rangle \\ &= \int \int L\,(\chi_0, \frac{1}{2} + it) \widehat{\psi}(t)\, b(x)\, |x|^{\frac{1}{2}+it}\, d^*x\, dt\,. \end{aligned} \tag{2.439}$$

Thus, from (2.427) and (2.428) we conclude, using arbitrary test functions b, that the Fourier transform of the distribution $L\,(\chi_0, 1/2 + it)\, \widehat{\psi}(t)$ actually vanishes, so that,

$$L\,(\chi_0, \frac{1}{2} + it)\, \widehat{\psi}(t) = 0\,. \tag{2.440}$$

To justify the above equality, we need to control the growth of the L-function in the variable t. One has

$$|L(\frac{1}{2} + it)| = O\,(|t|^N)\,. \tag{2.441}$$

In particular, since $L\left(\frac{1}{2} + it\right)$ is an analytic function of t, we see that it is a multiplier of the algebra $\mathcal{S}(\mathbb{R})$ of Schwartz functions in the variable t. Thus the product $L\,(\frac{1}{2} + it)\, \widehat{\psi}(t)$ is still a tempered distribution, and so is its Fourier transform. To say that the latter vanishes when tested on arbitrary functions which are smooth with compact support implies that it vanishes.

The above argument uses the hypothesis $\chi_0|_{C_{\mathbb{K},1}} \neq 1$. In the case $\chi_0|_{C_{\mathbb{K},1}} = 1$ we need to impose on the test function f used in (2.436) the condition $\int f\, dx = 0$, which means

$$\int b(x)\, |x|\, d^*x = 0\,. \tag{2.442}$$

But the space of functions $b(x)\,|x|^{1/2} \in C_c^\infty(\mathbb{R}_+^*)$ such that (2.442) holds is still dense in the Schwartz space $\mathcal{S}(\mathbb{R}_+^*)$.

To understand the equation (2.440), let us consider an equation for distributions $\alpha(t)$ of the form

$$\varphi(t)\,\alpha(t) = 0\,, \tag{2.443}$$

where we first work with distributions α on S^1 and we assume that $\varphi \in C^\infty(S^1)$ has finitely many zeros $x_i \in Z(\varphi)$, of finite order n_i. Let J be the ideal of $C^\infty(S^1)$ generated by φ. One has $\psi \in J \Leftrightarrow$ order of ψ at x_i is $\geq n_i$.

Thus the distributions $\delta_{x_i},\, \delta'_{x_i}, \ldots, \delta_{x_i}^{(n_i-1)}$ form a basis of the space of solutions of (2.443).

Now $\widehat{\psi}(t)$ is, for η orthogonal to Range $\mathfrak{E}$ and satisfying (2.429), a distribution with compact support, and $L\left(\chi_0, \frac{1}{2}+it\right)\widehat{\psi}(t) = 0$. Thus by the above argument we get that $\widehat{\psi}$ is a finite linear combination of the distributions $\delta_t^{(k)}$ with t such that $L\left(\chi_0, \frac{1}{2}+it\right) = 0$ and $k <$ order of the zero, $k < \frac{\delta-1}{2}$. The condition $k <$ order of the zero is necessary and sufficient to get the vanishing on the range of $\mathfrak{E}$. The condition $k < \frac{\delta-1}{2}$ is necessary and sufficient to ensure that ψ belongs to $L^2_{-\delta}$, i.e., that

$$\int (\log|x|)^{2k}\,(1+|\log|x||^2)^{-\delta/2}\,d^*x < \infty \tag{2.444}$$

which is $2k-\delta < -1$, i.e., $k < \frac{\delta-1}{2}$.

Conversely, let s be a zero of $L(\chi_0, s)$ and $k>0$ its order. By Lemma 2.52 and the finiteness and analyticity of Δ'_s (for $\Re s > 0$) we get

$$\left(\frac{\partial}{\partial s}\right)^a \Delta_s(f) = 0 \quad \forall\, f \in \mathcal{S}(\mathbb{A}_\mathbb{K})_0\,,\ a = 0,1,\ldots,k-1\,. \tag{2.445}$$

We can differentiate the equality of Lemma 2.52 and get

$$\left(\frac{\partial}{\partial s}\right)^a \Delta_s(f) = \int_{C_\mathbb{K}} \mathfrak{E}(f)(x)\,\chi_0(x)\,|x|^{s-1/2}\,(\log|x|)^a\,d^*x\,. \tag{2.446}$$

Thus η belongs to the orthogonal of Range $\mathfrak{E}$ and satisfies (2.429) iff it is a finite linear combination of functions of the form

$$\eta_{t,a}(x) = \chi_0(x)\,|x|^{it}\,(\log|x|)^a, \tag{2.447}$$

where

$$L\left(\chi_0, \frac{1}{2}+it\right) = 0\ , \quad a < \text{order of the zero},\ a < \frac{\delta-1}{2}\,. \tag{2.448}$$

The restriction to the subgroup $\mathbb{R}_+^*$ of $C_\mathbb{K}$ of the transpose of $\underline{\vartheta}_\mathrm{m}$ is thus given in the above basis by

$$\underline{\vartheta}_\mathrm{m}(\lambda)^t\,\eta_{t,a} = \sum_{b=0}^{a} C_a^b\,\lambda^{it}\,(\log(\lambda))^b\,\eta_{t,a-b}. \tag{2.449}$$

The multiplication operator by a function with bounded derivatives is a bounded operator in any Sobolev space, as one checks directly using the density in the orthogonal of Range $\mathfrak{E}$ of vectors satisfying (2.429), that is, if $L\left(\chi_0, \frac{1}{2}+is\right) \neq 0$, then is does not belong to the spectrum of $D^t_{\chi_0}$.

This determines the spectrum of the operator $D^t_{\chi_0}$ and hence of its transpose D_{χ_0} as indicated in Theorem 2.47 and ends the proof of Theorem 2.47.

Let us now prove corollary 2.48. Let us fix $h_0 \in \mathcal{S}(C_{\mathbb{K}})$ such that $\widehat{h}_0$ has compact support contained in $\{\chi_0\} \times \mathbb{R}$ and $\widehat{h}_0(\chi_0, s) = 1$ for s small.

Let then h_s be given by $h_s(g) = h_0(g)\,|g|^{is}$. The Fourier transform $\widehat{h}_s$ is then the translate of $\widehat{h}_0$, and one can choose h_0 such that

$$\sum_{n\in\mathbb{Z}} \widehat{h}_n(\chi_0\,, u) = 1 \quad u \in \mathbb{R}\,. \tag{2.450}$$

When $|s| \to \infty$, the dimension of the range of $\underline{\vartheta}_{\mathrm{m}}{}^t(h_s)$ is of the order of $\log|s|$ as is the number of zeros of the L-function in the translates of a fixed interval (cf. [**294**]).

Let $h \in \mathcal{S}(C_{\mathbb{K}})$. One has $\underline{\vartheta}_{\mathrm{m}}{}^t(h) = \sum_{n\in\mathbb{Z}} \underline{\vartheta}_{\mathrm{m}}{}^t(h * h_n)$. It follows then from the polynomial growth of the norm of $\underline{\vartheta}_{\mathrm{m}}{}^t(g)$ that the operator

$$\int h(g)\,\underline{\vartheta}_{\mathrm{m}}(g)^t\, d^*g \tag{2.451}$$

is of trace-class for any $h \in \mathcal{S}(C_{\mathbb{K}})$.

Moreover using the triangular form given by (2.449) we get its trace, and hence the trace of its transpose $\underline{\vartheta}_{\mathrm{m}}(h)$ as

$$\operatorname{Trace} \underline{\vartheta}_{\mathrm{m}}(h) = \sum_{\substack{L\left(\chi,\frac{1}{2}+\rho\right)=0 \\ \rho\in i\,\mathbb{R}}} \widehat{h}(\chi, \rho) \tag{2.452}$$

where the multiplicity is counted as in Theorem 2.47 and where the Fourier transform $\widehat{h}$ of h is defined by

$$\widehat{h}(\chi, \rho) = \int_{C_{\mathbb{K}}} h(u)\,\widetilde{\chi}(u)\,|u|^\rho\, d^*\,u\,. \tag{2.453}$$

10. A Lefschetz formula for Archimedean local factors

As we saw above, the noncommutative space $X_{\mathbb{K}}$ of adèle classes over a global field $\mathbb{K}$ provides both a spectral realization of zeros of L-functions and an interpretation of the explicit formulas of Riemann-Weil as a Lefschetz formula. The corresponding trace formula was proved in [**71**] in the semilocal case (finitely many places). It is natural then to ask whether a similar approach can be applied to the L-functions of motives, also in view of the fact that the results of [**102**], [**103**] showed that noncommutative geometry can be employed to describe properties of the fibers at Archimedean places (and at places of totally degenerate reduction) of arithmetic varieties.

In this section we consider the Archimedean factors of the Hasse-Weil L-function attached to a non-singular projective algebraic variety X defined over a number field $\mathbb{K}$. We obtain the real part of the logarithmic derivative of the Archimedean factors of the L-function $L(H^m, z)$ on the critical line as a trace formula for the action of a suitable Weil group on a complex manifold attached to the Archimedean place.

In §10.3 we make some preliminary calculations. First we show (Lemma 2.55) that for $\mathbb{K}_v = \mathbb{C}$ or $\mathbb{R}$, one obtains the logarithmic derivative of the imaginary part of the gamma functions $\Gamma_{\mathbb{K}_v}$ (2.461) on the critical line as the principal value on $\mathbb{K}_v^*$ of a distribution

$$\frac{|u|^{\frac{1}{2}+i\,s}}{|1-u|}.$$

We then show (Lemma 2.57) that the shift by $\min\{p,q\}$ in the argument of $\Gamma_{\mathbb{C}}$ appears when one considers the principal value on $\mathbb{C}^*$ of

$$\frac{u^{-p}\overline{u}^{-q}|u|_{\mathbb{C}}^{z}}{|1-u|_{\mathbb{C}}}.$$

This is sufficient to obtain in Theorem 2.58 of §10.4 the logarithmic derivative of the Archimedean local factor as a trace formula for the action of $\mathbb{C}^*$ on a space with base $\mathbb{C}$ and fiber $H^m(X,\mathbb{C})$ with the representation

$$\pi(H^m,u)\,\xi = u^{-p}\overline{u}^{-q}\xi \tag{2.454}$$

of $\mathbb{C}^*$ on the cohomology. The real case is more delicate and it is treated in §10.5. In this case, to obtain a similar result, one needs to consider a space that has base the quaternions $\mathbb{H}$ and fiber the cohomology. We then obtain the logarithmic derivative of the local factor as a trace formula for an action on this space of the Weil group, with representation

$$\pi(H^m,wj)\,\xi = \,i^{p+q}\,w^{-p}\overline{w}^{-q}F_\infty(\xi) \tag{2.455}$$

on the cohomology, with F_∞ the linear involution induced by complex conjugation. In §10.6 we describe the problem of a semilocal trace formula involving several Archimedean places. In §10.7 we address the problem of considering simultaneously Archimedean and non-Archimedean places and in particular the fact that in the usual approaches one is forced to make a choice of embeddings of $\mathbb{Q}_\ell$ in $\mathbb{C}$. We suggest that, at least in the case of a curve X, one may obtain a different approach by considering, as the fiber of the space one would like to construct for a semi-local Lefschetz trace formula, the adèle class space of the function field of the curve over the residue field k_v at the place v.

10.1. Archimedean local L-factors.

We saw in Chapter I §8.1.3 the definition of L-functions of algebraic varieties defined over a number field $\mathbb{K}$ as an Euler product (1.380) of local factors (1.381). Let us briefly recall this definition of the L-functions $L(H^m(X),s)$ of a smooth projective algebraic variety over a number field $\mathbb{K}$. It is written as an Euler product over finite and Archimedean places v of $\mathbb{K}$,

$$L(H^m(X),s) = \prod_v L_v(H^m(X),s). \tag{2.456}$$

The local L-factors of (2.456) at finite primes encode the action of the geometric Frobenius on the inertia invariants $H^m(\bar{X},\mathbb{Q}_\ell)^{I_v}$ of the étale cohomology in the form (cf. [**265**])

$$L_v(H^m(X),s) = \det\left(1 - Fr_v^* N(v)^{-s}|H^m(\bar{X},\mathbb{Q}_\ell)^{I_v}\right)^{-1}, \tag{2.457}$$

with N the norm map. Historically, the roots of the idea of motives introduced by Grothendieck can be found in the L-functions $L(X,s)$ of algebraic varieties, which can be written as an alternating product

$$L(X,s) = \prod_{i=0}^{n} L(H^i(X),s)^{(-1)^{i+1}}. \tag{2.458}$$

Typically, such a combined L-function as (2.458) tends to have infinitely many poles, corresponding to the zeros of the factors in the denominator, whereas in isolating a single factor $L(H^i(X),s)$ one avoids this problem and also gains the possibility of having a nice functional equation. It is convenient to think of such a factor $L(H^i(X),s)$ as the L-function, not of a variety, but of a pure motive of weight i, viewed as a direct summand of the variety X.

Serre showed in [**265**] that the local factors one needs to consider at the Archimedean primes, dictated by the expected form of the functional equation for the L-function $L(H^m(X),s)$, depend upon the Hodge structure

$$H^m(X_v(\mathbb{C})) = \oplus_{p+q=m} H^{p,\,q}(X_v(\mathbb{C})), \tag{2.459}$$

where $v : \mathbb{K} \hookrightarrow \mathbb{C}$ is the Archimedean place and $X_v(\mathbb{C})$ is the corresponding complex algebraic variety.

The Archimedean local factors have an explicit formula (cf. [**265**]) given in terms of gamma functions, with shifts in the argument and powers that depend on the Hodge numbers. Namely, one has a product of gamma functions according to the Hodge numbers $h^{p,q}$ of the form

$$L(H^*,s) = \begin{cases} \prod_{p,q} \Gamma_{\mathbb{C}}(s-\min(p,q))^{h^{p,q}} \\ \prod_{p<q} \Gamma_{\mathbb{C}}(s-p)^{h^{p,q}} \prod_p \Gamma_{\mathbb{R}}(s-p)^{h^{p,+}} \Gamma_{\mathbb{R}}(s-p+1)^{h^{p,-}} \end{cases} \tag{2.460}$$

where the two cases correspond, respectively, to the complex and the real embeddings, $\mathbb{K}_v = \mathbb{C}$ and $\mathbb{K}_v = \mathbb{R}$. Here $h^{p,\pm}$ is the dimension of the $\pm(-1)^p$-eigenspace of the involution on $H^{p,p}$ induced by the real structure and

$$\Gamma_{\mathbb{C}}(s) := (2\pi)^{-s}\Gamma(s), \quad \Gamma_{\mathbb{R}}(s) := 2^{-1/2}\pi^{-s/2}\Gamma(s/2). \tag{2.461}$$

It is clear that it is desirable to have a reformulation of the formulae for the local factors in such a way that the Archimedean and the non-Archimedean cases are treated as much as possible on equal footing. To this purpose, Deninger in [**116**], [**117**] expressed both (2.457) and (2.460) in the form of zeta-regularized infinite determinants. Here we propose a different approach. Namely, we reinterpret the Archimedean local factors as a Lefschetz trace formula over a suitable geometric space. The main idea is that we expect a global (or at least semi-local) Lefschetz trace formula to exist for the L-functions $L(H^m(X),s)$, over a noncommutative space which will be a noncommutative generalization of the pure motive $h^m(X)$ and should be obtained as an extension of $h^m(X)$ by a suitable modification of the adèle class space (by certain division algebras instead of the local fields $\mathbb{K}_v$ at the places v of bad reduction and at the real Archimedean place). Here we give some evidence for this idea, by giving the local formula for a single Archimedean place v. One sees already in the case of a single real place that the underlying geometric space for the trace formula is obtained by passing to the division algebra of quaternions.

10.2. Asymptotic form of the number of zeros of L-functions.

The properties of the L-functions of algebraic varieties defined over a number field $\mathbb{K}$ are mostly conjectural ([**265**]). In particular one expects from the functional equation ([**265**], C_9) that the zeros of $L(H^m(X), s)$ are located on the critical line $\Re(s) = \frac{1+m}{2}$. By analogy with the case of the Riemann zeta function discussed above in §3 one expects that the number of zeros

$$N(E) = \#\{\rho \,|\, L(H^m(X), \rho) = 0, \quad \text{and} \quad 0 < \Im(\rho) \leq E\}\,, \tag{2.462}$$

can be decomposed as the sum of an average part and an oscillatory part

$$N(E) = \langle N(E) \rangle + N_{\text{osc}}(E)\,, \tag{2.463}$$

where the average part is obtained following Lemma 2.1, using the Archimedean local factors $L_v(H^m(X), s)$, in the form

$$\langle N(E) \rangle = \frac{1}{\pi} \sum_{v|\infty} \Im \log L_v(H^m(X), \frac{1+m}{2} + iE)\,. \tag{2.464}$$

One can rewrite (2.464) in a more symmetric manner involving all non-trivial zeros, with

$$N_s(E) = \#\{\rho \,|\, L(H^m(X), \rho) = 0, \quad \text{and} \quad -E \leq \Im(\rho) \leq E\}\,, \tag{2.465}$$

and eliminating the ambiguity in the definition of the log, so that (2.464) becomes

$$\langle N_s(E) \rangle = \sum_{v|\infty} \frac{1}{\pi} \int_{-E}^{E} \frac{d}{ds} \Im \log L_v(H^m(X), \frac{1+m}{2} + is) ds\,. \tag{2.466}$$

We shall now show that the key ingredient of this formula can be expressed in the form of a Lefschetz contribution

$$\frac{1}{\pi} \frac{d}{ds} \Im \log L_v(H^m(X), \frac{1+m}{2} + is) = -\frac{1}{2\pi} \int'_{W_v} \frac{\mathrm{Tr}(\pi_v(H^m, u))|u|^z_{W_v}}{|1-u|_{\mathbb{H}_v}} d^*u \tag{2.467}$$

where $z = \frac{1+m}{2} + is$, the group W_v is a local Weil group which is a modulated group. It embeds in the algebra $\mathbb{H}_v$ which is the algebra $\mathbb{H}$ of quaternions for v a real place and the algebra $\mathbb{C}$ for a complex place. We shall describe below in §10.4 and 10.5 the natural representation $\pi_v(H^m, u)$ of W_v.

10.3. Weil form of logarithmic derivatives of local factors.

We begin by checking the formulae ([**294**], [**71**]) that will relate the Fourier transform of the local Lefschetz contribution (viewed as a distribution on the multiplicative group) with the derivative of the imaginary part of the logarithm of the Archimedean local factor $\Gamma_{\mathbb{K}_v}$, for $\mathbb{K}_v = \mathbb{R}$ (v a real Archimedean place) or $\mathbb{K}_v = \mathbb{C}$ (v a complex Archimedean place).
We use the notations of [**71**] for the principal values. We recall the following basic facts.

Lemma 2.55. *For $\mathbb{K}_v = \mathbb{R}$ or $\mathbb{K}_v = \mathbb{C}$ and for s real, one has*

$$\int'_{\mathbb{K}_v^*} \frac{|u|^{\frac{1}{2}+is}}{|1-u|} d^*u = -2 \frac{d}{ds} \Im \log \Gamma_{\mathbb{K}_v} \left(\frac{1}{2} + is\right)\,. \tag{2.468}$$

PROOF. First notice that, for s real, one has

$$2i\,\Im \log \Gamma_{\mathbb{K}_v}\left(\frac{1}{2}+is\right) = \log \Gamma_{\mathbb{K}_v}\left(\frac{1}{2}+is\right) - \log \Gamma_{\mathbb{K}_v}\left(\frac{1}{2}-is\right), \tag{2.469}$$

since $\Gamma_{\mathbb{K}_v}$ defined in (2.461) is a "real" function, i.e. it fulfills $f(\overline{z}) = \overline{f(z)}$.
Thus, one can rewrite the equality that needs to be proved in the form

$$\int'_{\mathbb{K}_v^*} \frac{|u|^{\frac{1}{2}+is}}{|1-u|} d^*u = -\left(\frac{\Gamma'_{\mathbb{K}_v}}{\Gamma_{\mathbb{K}_v}}\left(\frac{1}{2}+is\right) + \frac{\Gamma'_{\mathbb{K}_v}}{\Gamma_{\mathbb{K}_v}}\left(\frac{1}{2}-is\right)\right). \tag{2.470}$$

The i in is brings up the real part of $\frac{\Gamma'_{\mathbb{K}_v}}{\Gamma_{\mathbb{K}_v}}$ which is again a "real" function.
Notice also that (2.470) should hold without the restriction $s \in \mathbb{R}$, since both sides are analytic functions of $s \in \mathbb{C}$.
Let us first take $\mathbb{K}_v = \mathbb{R}$. In this case we have

$$\Gamma_{\mathbb{R}}(x) = 2^{-1/2}\pi^{-x/2}\Gamma(x/2), \tag{2.471}$$

hence the equality takes the explicit form in terms of the usual Γ-function

$$\int'_{\mathbb{R}^*} |u|^{is}\frac{|u|^{1/2}}{|1-u|} d^*u = \log \pi - \frac{1}{2}\left(\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}+i\frac{s}{2}\right) + \frac{\Gamma'}{\Gamma}\left(\frac{1}{4}-i\frac{s}{2}\right)\right). \tag{2.472}$$

which follows from (2.130). Let us then take $\mathbb{K}_v = \mathbb{C}$. In this case we have

$$\Gamma_{\mathbb{C}}(x) = (2\,\pi)^{-x}\Gamma(x), \tag{2.473}$$

and again, in terms of the usual Γ-function, the equality takes the form

$$\int'_{\mathbb{C}^*} |u|^{is}\frac{|u|^{1/2}}{|1-u|} d^*u = 2\,\log 2\pi - \left(\frac{\Gamma'}{\Gamma}\left(\frac{1}{2}+is\right) + \frac{\Gamma'}{\Gamma}\left(\frac{1}{2}-is\right)\right). \tag{2.474}$$

This can be extracted from [**71**] (Appendix 2) and will be proved as a special case of Lemma 2.57 below. ∎

The next step consists in obtaining a similar formula for

$$\frac{d}{ds}\Im \log \Gamma_{\mathbb{K}_v}\left(\frac{1}{2}+\frac{|n|}{2}+is\right), \tag{2.475}$$

where $n \in \mathbb{Z}$ is an integer. Notice that the function $\Gamma_{\mathbb{K}_v}(\frac{|n|}{2}+z)$ is still real. We take $\mathbb{K}_v = \mathbb{C}$ and we let

$$f_0(\nu) = \inf(\nu^{1/2}, \nu^{-1/2}), \quad \forall \nu \in \mathbb{R}_+^*. \tag{2.476}$$

With this notation one gets,

LEMMA 2.56. *For $n \in \mathbb{Z}$ and $\rho \in \mathbb{R}_+^*$ with $\rho \neq 1$, one has*

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{in\theta}}{|1-e^{i\theta}\rho|^2}\, d\theta = \frac{f_0(\nu)^{|n|}}{|1-\nu|}, \qquad \nu = \rho^2. \tag{2.477}$$

PROOF. Let us first consider the case $\rho < 1$. Then we have

$$\frac{1-\rho^2}{|1-e^{i\theta}\rho|^2} = \frac{1}{1-\rho e^{i\theta}} + \frac{1}{1-\rho e^{-i\theta}} - 1, \tag{2.478}$$

which gives the Fourier coefficients $\rho^{|n|}$ for the function on the left hand side. This gives the equality in the case $\rho < 1$, since $f_0(\nu) = \rho$ in that case. One then checks that both sides of (2.477) fulfill

$$f(\rho^{-1}) = \rho^2 f(\rho)\,,$$

which gives the desired equality for $\rho > 1$. $\square$

We then obtain the following result:

LEMMA 2.57. *Let $\mathbb{K}_v = \mathbb{C}$ and suppose given $p, q \in \mathbb{N}$ with $m = p + q$. For $z = \frac{1+m}{2} + is$, with $s \in \mathbb{R}$, one has*

$$\int'_{\mathbb{C}^*} \frac{u^{-p}\overline{u}^{-q}|u|_{\mathbb{C}}^z}{|1-u|_{\mathbb{C}}} d^*u = -2\,\frac{d}{ds}\Im \log \Gamma_{\mathbb{C}}\left(z - \min(p,q)\right)\,. \tag{2.479}$$

PROOF. Let $n = p - q$. One has $\min(p,q) = \frac{m}{2} - \frac{|n|}{2}$. Since $|u|_{\mathbb{C}} = u\overline{u}$, one has $u^{-p}\overline{u}^{-q} = e^{-in\theta}\,|u|_{\mathbb{C}}^{\frac{-m}{2}}$, where θ is the argument of u. One can then rewrite the desired equality in the form

$$\int'_{\mathbb{C}^*} \frac{e^{-in\theta}\,|u|_{\mathbb{C}}^{\frac{1}{2}+is}}{|1-u|_{\mathbb{C}}} d^*u = -2\,\frac{d}{ds}\Im \log \Gamma_{\mathbb{C}}\left(\frac{1}{2} + is + \frac{|n|}{2}\right)\,. \tag{2.480}$$

Recall now that $\int'$ coincides with the Weil principal value Pfw which is defined in (2.293) and (2.294) so that

$$Pfw \int_{\mathbb{C}^*} \varphi(u)\, d^*u = PF_0 \int_{\mathbb{R}_+^*} \psi(\nu)\, d^*\nu\,, \tag{2.481}$$

where $\psi(\nu) = \int_{|u|_{\mathbb{C}}=\nu} \varphi(u)\, d_\nu u$ is obtained by integrating φ over the fibers, while one has

$$\begin{aligned} &PF_0 \textstyle\int \psi(\nu) d^*\nu = \\ &\qquad 2\log(2\pi)\, c + \lim_{t\to\infty} \left(\textstyle\int (1 - f_0^{2t})\, \psi(\nu)\, d^*\nu - 2c\log t\right)\,. \end{aligned} \tag{2.482}$$

where $\psi - c\, f_1^{-1}$ is integrable on $\mathbb{R}_+^*$, with $f_1 = f_0^{-1} - f_0$. In our case we have

$$\varphi(u) = \frac{e^{-in\theta}\,|u|_{\mathbb{C}}^{\frac{1}{2}+i\,s}}{|1-u|_{\mathbb{C}}}.$$

By Lemma 2.56, integration over the fibers gives

$$\psi(\nu) = \frac{\nu^{\frac{1}{2}+i\,s}\, f_0(\nu)^{|n|}}{|1-\nu|}\,.$$

We have $c = 1$ independently of s and n. One has the identity (2.335),

$$PF_0 \int f_0\, f_1^{-1} d^*\nu = 2(\log 2\pi + \gamma)\,.$$

This allows one to check the equality for $n = 1$ and $s = 0$. Indeed, in that case one gets $\psi(\nu) = f_0\, f_1^{-1}$, while

$$-2\left(\frac{\Gamma'_{\mathbb{C}}}{\Gamma_{\mathbb{C}}}\left(\frac{1+n}{2}\right)\right) = -2(-\log(2\pi) - \gamma)\,.$$

Having checked the result for some value of n and s, while c is independent of both, one gets convergence of the integral for the variations and one can now use any regularization to compare other values of

$$\int_{\mathbb{R}_+^*} \frac{\nu^{\frac{1}{2}+i\,s}\, f_0(\nu)^{|n|}}{|1-\nu|}\, d^*\nu\,. \tag{2.483}$$

We write the integral as $\int_0^1 + \int_1^\infty$ and use the minimal substraction as regularization i.e. the substraction of the pole part in ε after replacing the denominator $|1-\nu|$ by $|1-\nu|^{1-\varepsilon}$. The first integral gives

$$\int_0^1 \nu^{\left(\frac{1}{2}+is+\frac{|n|}{2}\right)}(1-\nu)^{-1+\varepsilon}\frac{d\nu}{\nu} = B\left(\frac{1}{2}+is+\frac{|n|}{2},\varepsilon\right) = \frac{\Gamma\left(\frac{1}{2}+is+\frac{|n|}{2}\right)\Gamma(\varepsilon)}{\Gamma\left(\frac{1}{2}+is+\frac{|n|}{2}+\varepsilon\right)}\,.$$

The residue at $\varepsilon = 0$ is equal to one and the finite part gives

$$-\frac{\Gamma'}{\Gamma}\left(\frac{1}{2}+is+\frac{|n|}{2}\right)-\gamma\,. \tag{2.484}$$

The other integral $\int_1^\infty$ gives the complex conjugate,

$$-\frac{\Gamma'}{\Gamma}\left(\frac{1}{2}-is+\frac{|n|}{2}\right)-\gamma\,. \tag{2.485}$$

Thus, having checked the additive constant term, we get

$$\int_{\mathbb{C}^*}' \frac{e^{-in\theta}\,|u|_{\mathbb{C}}^{\frac{1}{2}+is}}{|1-u|_{\mathbb{C}}}d^*u = -2\,\frac{d}{ds}\Im\log\Gamma_{\mathbb{C}}\left(\frac{1}{2}+i\,s+\frac{|n|}{2}\right)\,, \tag{2.486}$$

and the required equality follows. □

10.4. Lefschetz formula for complex places.

We now look at a complex Archimedean place with the local L-factor $L_{\mathbb{C}}(H^m, z)$. We let $\pi(H^m, u)$ be the canonical representation ([**109**]) of $\mathbb{C}^*$ on $H^m = \oplus H^{p,q}$,

$$\pi(H^m,u)\,\xi = u^{-p}\overline{u}^{-q}\xi\,,\quad \forall \xi\in H^{p,q}\,,\ \ \forall u\in\mathbb{C}^*\,. \tag{2.487}$$

We obtain the following formula for the Archimedean local factor at a complex place.

THEOREM 2.58. *Let $\mathbb{K}_v = \mathbb{C}$. For $m\in\mathbb{N}$, let $\pi(H^m,u)$ be the canonical representation of $\mathbb{C}^*$ on H^m. Then, for $z = \frac{1+m}{2} + is$ with $s\in\mathbb{R}$, we have*

$$\int_{\mathbb{C}^*}' \frac{\mathrm{Tr}(\pi(H^m,u))|u|_{\mathbb{C}}^z}{|1-u|_{\mathbb{C}}}d^*u = -2\,\frac{d}{ds}\Im\log L_{\mathbb{C}}(H^m,z)\,. \tag{2.488}$$

PROOF. This follows directly from Lemma 2.57 and the formula (2.460) (cf. [**265**]) that expresses $L_{\mathbb{C}}(H^m,z)$ as a product of powers of the $\Gamma_{\mathbb{C}}(z-\min(p,q))$ and hence the logarithmic derivative as a sum

$$\frac{d}{ds}\Im\log L_{\mathbb{C}}(H^m,z) = \sum_{p+q=m} h^{p,q}\,\frac{d}{ds}\Im\log\Gamma_{\mathbb{C}}(z-\min(p,q))$$

where $h^{p,q}$ is the dimension of $H^{p,q}$. □

10.5. Lefschetz formula for real places.

We now look at a real Archimedean place with the local L-factor $L_{\mathbb{R}}(H^m, z)$. We let W be the Weil group [**294**] which, in this case, is the normalizer of $\mathbb{C}^*$ in $\mathbb{H}^*$ where $\mathbb{H}$ here denotes the division algebra of quaternions $\mathbb{H} = \mathbb{C} + \mathbb{C}j$, with $j^2 = -1$ and $jwj^{-1} = \overline{w}$, $\forall w \in \mathbb{C}$. Elements of W are of the form

$$u = wj^{\epsilon}\,, \quad w \in \mathbb{C}^*\,, \quad \epsilon \in \{0, 1\}\,,$$

where we use the notation $j^0 = 1$, $j^1 = j$.
We let $\pi(H^m, u)$ be the canonical representation ([**109**]) of the Weil group in $H^m = \oplus H^{p,q}$. The subgroup $\mathbb{C}^* \subset W$ acts as above, while elements in $\mathbb{C}^* j$ act by

$$\pi(H^m, wj)\,\xi = i^{p+q}\, w^{-p}\overline{w}^{\,-q} F_\infty(\xi)\,, \quad \forall \xi \in H^{p,q}\,. \tag{2.489}$$

Here F_∞ is the linear involution associated to the geometric action of complex conjugation (once translated to the cohomology with complex coefficients) as in Serre ([**265**]).
One has to check that π is a representation, and in particular that $\pi(H^m, j)^2 = \pi(H^m, -1)$. This follows for $\xi \in H^{p,q}$ from

$$\pi(H^m, j)^2 \xi = (-1)^{p+q} F_\infty^2(\xi) = (-1)^{p+q}\xi\,.$$

One checks in the same way that $\pi(H^m, j)\,\pi(H^m, w) = \pi(H^m, \overline{w})\,\pi(H^m, j)$.
We now investigate the integral

$$\int_W' \frac{\operatorname{Tr}(\pi(H^m, u))|u|_{\mathbb{H}}^z}{|1-u|_{\mathbb{H}}} d^*u\,, \tag{2.490}$$

in which we follow the conventions of Weil [**294**]. The reduced norm of a quaternion is

$$|q|_{\mathbb{H}} = a\bar{a} + b\bar{b} = |a|^2 + |b|^2\,, \quad q = a + bj\,, \quad \forall a\,, b \in \mathbb{C} \tag{2.491}$$

so that for the natural inclusion $\mathbb{C} \subset \mathbb{H}$ one has

$$|z|_{\mathbb{H}} = |z|_{\mathbb{C}}\,, \quad \forall z \in \mathbb{C} \subset \mathbb{H}\,. \tag{2.492}$$

Thus, for $u \in W$, the module $|u|_{\mathbb{H}}$ is the natural module $|u|_W$ of the Weil group. The finite value $\int'$ is defined by integration over the fibers of the modulated group $W \to \mathbb{R}_+^*$ using (2.481) and (2.482).

Let us start with the case m odd, since then the subtle term involving $\Gamma_{\mathbb{R}}$ does not enter. Then the action of W on $H^{p,q} \oplus H^{q,p}$ will give the same trace as in the complex case for elements $u = wj^0 = w \in \mathbb{C}^*$, but it will give zero trace to any element $u = wj^1 = wj$. In fact, in this case the action is given by an off-diagonal matrix since it exchanges the subspaces $H^{p,q}$ and $H^{q,p}$. Thus the function $\operatorname{Tr}(\pi(H^m, u))$ vanishes except on $\mathbb{C}^* \subset W$ and it agrees there with what it would be forgetting about the real structure. Integration over the fibers of the modulated group $W \to \mathbb{R}_+^*$ using (2.481) and (2.482) gives an overall factor of $\frac{1}{2}$ in the above expression, since when one integrates over the fibers of the module map $u \mapsto |u|_{\mathbb{H}}$ from W to $\mathbb{R}_+^*$ the only contribution will come from the $u = wj^0 = w$. This will give the same answer as for the complex case, except for an overall factor $\frac{1}{2}$ due to the normalization of the fiber measure as a probability measure. Thus, this gives the right result when m is odd, since in that case the local factor $L_{\mathbb{R}}(H^m, z)$ is really the square root of what it would be when viewed as complex. Indeed, it reduces to

$$\prod_{p<q} \Gamma_{\mathbb{C}}(s-p)^{h^{p,q}}$$

which is the square root of

$$\prod_{p,q} \Gamma_{\mathbb{C}}(z-\min(p,q))^{h^{p,q}}$$

When $m = 2p$ is even and $h^{p+} = h^{p-}$, the same argument does apply irrespectively of the detailed definition of the representation of W. Indeed, elements of W of the form $u = wj$ act by

$$\pi(H^m, wj)\,\xi = (-1)^p\,|w|^{-2p}\,F_\infty(\xi)\,, \qquad \forall \xi \in H^{p,p}\,, \tag{2.493}$$

and have pairs of eigenvalues of opposite signs and vanishing trace. One thus gets the required result using (2.460) and the duplication formula (2.152), i.e.

$$\Gamma_{\mathbb{R}}(z)\,\Gamma_{\mathbb{R}}(z+1) = \Gamma_{\mathbb{C}}(z)\,. \tag{2.494}$$

When $h^{p+} \neq h^{p-}$, we use the detailed definition of the action of j. Notice that the computation will not be reducible to the previous ones, since one uses the quaternions to evaluate

$$|1-u|_{\mathbb{H}} = 1 + |w|^2\,, \quad u = wj\,, \tag{2.495}$$

The action of j is $(-1)^p$ times the geometric action of complex conjugation (once translated to the cohomology with complex coefficients) on the space $H^{p,p}$.
We compute in the following lemma the relevant integral.

LEMMA 2.59. *For $z = \frac{1}{2} + is$ with $s \in \mathbb{R}$, one has*

$$\int_{\mathbb{R}_+^*} \frac{u^z}{1+u} d^*u = -2\frac{d}{ds}\Im \log(\Gamma_{\mathbb{R}}(z)/\Gamma_{\mathbb{R}}(z+1))\,. \tag{2.496}$$

PROOF. First for $z \in \mathbb{C}$ with positive real part, one has

$$\int_0^1 \frac{u^z}{1+u} d^*u = \frac{1}{2}\left(\frac{\Gamma'}{\Gamma}\left(\frac{z+1}{2}\right) - \frac{\Gamma'}{\Gamma}\left(\frac{z}{2}\right)\right)\,. \tag{2.497}$$

Indeed, both sides fulfill $f(z) + f(z+1) = \frac{1}{z}$ and $f(z) \to 0$ when $z \to +\infty$.
Let us prove (2.496). The factors in $\pi^{-z/2}$ in (2.471) for $\Gamma_{\mathbb{R}}(z)$ do not contribute to the right hand side, which one can replace by

$$-\frac{1}{2}\left(\frac{\Gamma'}{\Gamma}\left(\frac{1}{4}+i\frac{s}{2}\right) + \frac{\Gamma'}{\Gamma}\left(\frac{1}{4}-i\frac{s}{2}\right)\right) + \frac{1}{2}\left(\frac{\Gamma'}{\Gamma}\left(\frac{3}{4}+i\frac{s}{2}\right) + \frac{\Gamma'}{\Gamma}\left(\frac{3}{4}-i\frac{s}{2}\right)\right)\,.$$

The equality (2.496) follows from (2.497), using $s \to -s$ and the symmetry of the integral to integrate from 1 to ∞. $\square$

We then obtain the following formula for the Archimedean local factor at a real place.

THEOREM 2.60. *Let $\mathbb{K}_v = \mathbb{R}$. For $m \in \mathbb{N}$, let $\pi(H^m, u)$ be the representation of the Weil group W on H^m described above. Then, for $z = \frac{1+m}{2} + is$ with $s \in \mathbb{R}$, we have*

$$\int_W' \frac{\mathrm{Tr}(\pi(H^m, u))|u|_{\mathbb{H}}^z}{|1-u|_{\mathbb{H}}} d^*u = -2\,\frac{d}{ds}\Im \log L_{\mathbb{R}}(H^m, z)\,. \tag{2.498}$$

PROOF. The discussion above shows that we are reduced to the case of even $m = 2p$ and the subspace $H^{p,p}$. Let then $k = h^{p+} - h^{p-}$. When $k = 0$ the discussion above already gives the result. When $k \neq 0$ the left hand side gets an additional term from the trace of elements of the form $u = wj \in W$. Their contribution is given by

$$\frac{k}{2}\int_{\mathbb{R}_+^*} \frac{v^{z-p}}{1+v} d^*v,$$

where the factor $\frac{1}{2}$ comes from the normalization (as a probability measure) of the fiber measure of the module $W \mapsto \mathbb{R}_+^*$. The term in v^{-p} comes from the action of the scalar part $w \in \mathbb{C}^*$ with $w\overline{w} = v$ in $u = wj$. The denominator comes from the equality

$$|1-u|_{\mathbb{H}} = 1 + |w|^2\,, \quad u = wj\,.$$

The other side changes in the same way using (2.496). Indeed, one has

$$h^{p+}\,\log\Gamma_{\mathbb{R}}(z-p) + h^{p-}\,\log\Gamma_{\mathbb{R}}(z-p+1) =$$

$$\frac{h^{p+}+h^{p-}}{2}\,\log(\Gamma_{\mathbb{R}}(z-p)\Gamma_{\mathbb{R}}(z-p+1)) + \frac{k}{2}\,\log(\Gamma_{\mathbb{R}}(z-p)/\Gamma_{\mathbb{R}}(z-p+1))$$

$$= \frac{h^{p+}+h^{p-}}{2}\,\log\Gamma_{\mathbb{C}}(z-p) + \frac{k}{2}\,\log(\Gamma_{\mathbb{R}}(z-p)/\Gamma_{\mathbb{R}}(z-p+1))\,.$$

This completes the proof using lemma 2.59. ∎

The space which gives the formula of Theorem 2.60 as a Lefschetz local contribution is obtained by taking the quaternions $\mathbb{H}$ as the base, but endowing them with an additional structure, namely their complex structure when viewed as a right vector space over $\mathbb{C}$. We will use the Atiyah-Bott Lefschetz formula (cf. [**9**]) applied to the $\bar\partial$-complex, which generates the crucial term of theorem 2.60, that is,

$$\frac{1}{|1-u|_{\mathbb{H}}}\,, \tag{2.499}$$

with the reduced norm $|u|_{\mathbb{H}}$ as above, while a more naive approach without the use of an elliptic complex would involve the square of the reduced norm. The Atiyah-Bott Lefschetz formula involves a numerator $\chi(u)$, which yields

$$\frac{\chi(u)}{|1-u|_{\mathbb{H}}^2} = \frac{1}{|1-u|_{\mathbb{H}}}\,. \tag{2.500}$$

In fact, the use of the $\bar\partial$-complex brings in the powers $\wedge^j T_{\mathbb{C}}(\mathbb{H})$ of the complex tangent space $T_{\mathbb{C}}(\mathbb{H})$ of the complex manifold $\mathbb{H}$ and the alternating sum

$$\chi(u) = \sum_j (-1)^j\,\chi_j(u) = \sum_j (-1)^j\,\mathrm{trace}(\wedge^j(u)). \tag{2.501}$$

This gives the determinant of the quaternion $1 - u \in \mathbb{H}$ viewed as a two by two complex matrix (even if some complex conjugates are involved) and

$$\chi(u) = |1-u|_{\mathbb{H}}. \tag{2.502}$$

10.6. The question of the spectral realization.

A more general problem will be to obtain for Archimedean local factors a trace formula analogous to the semilocal case of Theorem 4 of [**71**]. This means that one considers a number field K, with S be the set of all Archimedean places, and X a non-singular projective variety over $\mathbb{K}$. One would like to obtain the real part of the logarithmic derivative of the full Archimedean factor $L_S(H^m, z) = \prod_{v \in S} L_{\mathbb{K}_v}(H^m, z)$ on the critical line as a trace formula for the action of a suitable Weil group on a complex space.
For a single place, one can work with a vector bundle over a base space, that is, $B = \mathbb{C}$ for v complex and $B = \mathbb{H}$ for v real, with fiber the $\mathbb{Z}$-graded vector space $E^{(m)}$ given by the cohomology $H^m(X_v, \mathbb{C})$, with X_v the compact Kähler manifold determined by the embedding $v : \mathbb{K} \hookrightarrow \mathbb{C}$. One expects a trace formula of the following sort, analogous to the semilocal case of [**71**], modelled on the trace formula of [**157**].

CONJECTURE 2.61. *Let $h \in \mathcal{S}(\mathbb{R}_+^*)$ with compact support be viewed as an element of $\mathcal{S}(W)$ by composition with the module. Then, for $\Lambda \to \infty$, one has*

$$\operatorname{Tr}(R_\Lambda\, \vartheta_{\mathrm{a}}(h)) = \tag{2.503}$$
$$2h(1)B_m \log \Lambda + \sum_{v \in S} \int_{W_v}' \frac{h(u^{-1}) \operatorname{Tr}(\pi_v(H^m, u))}{|1-u|_{\mathbb{H}_v}}\, d^*u + o(1)$$

where B_m is the m-th Betti number of X, and $\int'$ is uniquely determined by the pairing with the unique distribution on $\mathbb{K}_v$ which agrees with $\frac{du}{|1-u|}$ for $u \neq 1$ and whose Fourier transform relative to α_v vanishes at 1.

For example, the above formula can be proved in the simplest case of a single complex place. The bundle $E = \oplus_m E^{(m)}$ comes endowed with a representation of $\mathbb{C}^*$, by

$$\lambda : (z, \omega) \mapsto (\lambda z, \lambda^{-p} \bar{\lambda}^{-q} \omega), \tag{2.504}$$

for $\omega \in H^{p,q}(X_v, \mathbb{C})$. We then let $\mathcal{H} = L^2(B, E^{(m)})$ be the Hilbert space of L^2-sections of $E^{(m)}$ (for its Hermitian metric of trivial bundle). The action of $W = \mathbb{C}^*$ on $\mathcal{H}$ is given by

$$(\vartheta_{\mathrm{a}}(\lambda)\xi)(b) = \lambda^p \bar{\lambda}^q \xi(\lambda^{-1} b)\,, \quad \forall \xi \in L^2(B, H^{p,q})\,, \tag{2.505}$$

where we identify sections of $E^{(m)}$ with $H^m(X_v, \mathbb{C})$-valued functions on B.
One can use a cutoff as in [**71**], by taking orthogonal projection P_Λ onto the subspace

$$P_\Lambda = \{\xi \in L^2(B, E^{(m)})\,;\ \xi(b) = 0\,,\ \forall b \in B\,,\ |b|_{\mathbb{C}} > \Lambda\}\,. \tag{2.506}$$

Thus, P_Λ is the multiplication operator by the function ρ_Λ, where $\rho_\Lambda(b) = 1$ if $|b|_{\mathbb{C}} \leq \Lambda$, and $\rho_\Lambda(b) = 0$ for $|b|_{\mathbb{C}} > \Lambda$. This gives an infrared cutoff and to get an ultraviolet cutoff we use $\widehat{P}_\Lambda = F P_\Lambda F^{-1}$, where F is the Fourier transform which depends upon the choice of the basic character α_v for the place v. We let

$$R_\Lambda = \widehat{P}_\Lambda\, P_\Lambda\,.$$

PROPOSITION 2.62. *For the set of places S consisting of a single complex place the trace formula* (2.503) *holds.*

PROOF. Both sides of the formula are additive functions of the representation of $\mathbb{C}^*$ in $H^m(X_v, \mathbb{C})$. We can thus assume that this representation corresponds to a one-dimensional $H^{p,q}$. Let then $h_1(\lambda) = \lambda^p \bar{\lambda}^q h(|\lambda|)$; one has

$$h_1(u^{-1}) = h(|u|^{-1}) \operatorname{Tr}(\pi_v(H^m, u)), \quad \forall u \in \mathbb{C}^*$$

while by (2.505) we get that $\vartheta_{\mathrm{a}}(h)$ is the same operator as $U(h_1)$ in the notations of [**71**]. Thus, applying Theorem 4 of [**71**] to h_1 gives the desired result. □

In discussing here the case of a single complex place v, we have taken the trivial bundle over $\mathbb{C}$ with fiber the cohomology $H^m(X_v, \mathbb{C})$. Already in order to treat the case of several complex places, one needs to use the following result ([**264**], proof of Proposition 12): the integers $h^{p,q}(X_v)$ are independent of the Archimedean place $v \in S$. This suggests that it is in fact more convenient to think of the fiber $E^{(m)}$ as the motive $h^m(X)$ rather than its realization. This leads naturally to the further question of a semi-local trace formula where the finite set of places S involves both Archimedean and non-Archimedean places. We discuss in section 10.7 below, in the case of curves, how one can think of a replacement for the ℓ-adic cohomology at the non-Archimedean places, using noncommutative geometry.

REMARK 2.63. The representation ϑ_{a} of Conjecture 2.61 is not unitary but, as in (2.52), the product

$$u \mapsto |u|_W^{-(1+m)/2} \vartheta_{\mathrm{a}}(u) \tag{2.507}$$

should be unitary. In particular let N_E be the spectral projection, for the scaling action, associated to the interval $[-E, E]$ in the dual group $\mathbb{R}$ of $\mathbb{R}_+^*$ as in Lemma 2.4. It is given by

$$N_E = \vartheta_{\mathrm{a}}(h_E^{(m)}), \quad \text{with} \quad h_E^{(m)}(u) = |u|_W^{-(1+m)/2} \frac{1}{2\pi} \int_{-E}^{E} |u|_W^{is} ds \,. \tag{2.508}$$

Applying the conjectured formula (2.503) to the function $h_E^{(m)}$ the left hand side gives the counting of quantum states, i.e. the analogue of (2.69),

$$\operatorname{Tr}(R_\Lambda N_E)\,. \tag{2.509}$$

The first term in the right hand side of (2.503) gives the analogue of (2.46), i.e. the contribution of the regular representation of the scaling group. Finally, using (2.467) and (2.464), the last terms of (2.503) combine to give

$$\sum_{v|\infty} \int'_{W_v} \frac{h_E^{(m)}(u^{-1}) \operatorname{Tr}(\pi_v(H^m, u))}{|1-u|_{\mathbb{H}_v}} d^*u$$
$$= \sum_{v|\infty} \frac{1}{2\pi} \int_{-E}^{E} \int'_{W_v} \frac{\operatorname{Tr}(\pi_v(H^m, u))|u|_{W_v}^{\frac{1+m}{2}+is}}{|1-u|_{\mathbb{H}_v}} d^*u\, ds$$
$$- \sum_{v|\infty} \frac{1}{\pi} \int_{-E}^{E} \frac{d}{ds} \Im \log L_v(H^m(X), \frac{1+m}{2} + is) ds = -\langle N_s(E) \rangle$$

which shows that one should expect that the zeros of the L-function appear as an absorption spectrum. Unlike in the case of the Riemann zeta function this remains conjectural in the above case of L-functions of arithmetic varieties.

10.7. Local factors for curves.

In the previous part of this section, we showed how to write the Archimedean local factors of the L-function of an algebraic variety X in the form of a Lefschetz trace formula. Eventually, one would like to obtain a semi-local formula, like the one conjectured above, not just for the Archimedean places, but for the full L-function. We are only making some rather speculative remarks at this point, and we simply want to illustrate the type of construction one expects should give the geometric side of such a Lefschetz trace formula. We just discuss the case where X is a curve for simplicity. In this case we can concentrate on the L-function for $H^1(X)$.
Let X be a curve over a number field $\mathbb{K}$. The local Euler factor at a place v has the following description (cf. [**265**]):

$$L_v(H^1(X),s) = \det\left(1 - Fr_v^* N(v)^{-s} | H^1(\bar{X},\mathbb{Q}_\ell)^{I_v}\right)^{-1}. \tag{2.510}$$

Here Fr_v^* is the geometric Frobenius acting on ℓ-adic cohomology of $\bar{X} = X \otimes \mathrm{Spec}(\bar{\mathbb{K}})$, with $\bar{\mathbb{K}}$ an algebraic closure and ℓ a prime with $(\ell,q)=1$, where q is the cardinality of the residue field k_v at v. We denote by N the norm map. The determinant is evaluated on the inertia invariants $H^1(\bar{X},\mathbb{Q}_\ell)^{I_v}$ at v (this is all of $H^1(\bar{X},\mathbb{Q}_\ell)$ when v is a place of good reduction). The L-function then has the Euler product description

$$L(H^1(X),s) = \prod_v L_v(H^1(X),s),$$

where for v a non-Archimedean place the local factor is of the form (2.510) and if v is a complex or real Archimedean place then the local factor is given by the corresponding Γ-factor as discussed in the previous part of this section.
Usually, in using the expression (2.510) for the local factors, one makes use of a choice of an embedding of $\mathbb{Q}_\ell$ in $\mathbb{C}$. In our setting, if one wants to obtain a semi-local trace formula for the L-function $L(H^1(X),s)$, one needs a geometric construction which does not depend on such choices and treats the Archimedean and non-Archimedean places on equal footing. One expects that a geometric space on which the geometric side of the desired Lefschetz trace formula will concentrate should be obtained as a fibration, where the base space should be a noncommutative space obtained from the adèle class space of the number field $\mathbb{K}$, modified by considering, at the places of bad reduction and at the real Archimedean places, suitable division algebras over the local field. The fiber should also be a noncommutative space in which the special fiber embeds (at least at the places of good reduction). This will have the effect of replacing the use of the ℓ-adic cohomology and the need for a choice of an embedding of $\mathbb{Q}_\ell$ in $\mathbb{C}$.
In the case of a curve X, one can obtain this by embedding the special fiber X_v at a place v, which is a curve over the residue field k_v of cardinality q, in the adèle class space of the function field of X_v. Indeed, by Theorem 2.36 applied to the global field $\mathbb{K}$ of functions on X_v, one obtains the local factor $L_v(H^1(X),s)$ directly over $\mathbb{C}$. Thus, at least in the case of curves, the adèle class spaces of the global fields of functions on the curves X_v should be essential ingredients in the construction. This would then make it possible to work everywhere with a cohomology defined over $\mathbb{C}$, in the form of cyclic (co)homology.

10.8. Analogy with dimensional regularization.

We shall end this chapter by a brief discussion of the relation of the above computation of the Archimedean local L-factors with the set-up of Chapter 1 for renormalization using equisingular connections.

The natural framework for the above computations of L-factors is that of *real mixed Hodge structure* [**107**] which was recalled in Chapter 1 §8.4. A real mixed Hodge structure is a complex vector space V endowed with three opposite filtrations $(W_\bullet, F^\bullet, \bar{F}^\bullet)$ and a complex conjugation which respects W and interchanges F and $\bar{F}$. By [**107**] Proposition 1.2, this structure can be described equivalently (using the functor $W \mapsto \mathrm{Gr}^W V$) as a bigraded complex vector space

$$V = \oplus V^{p,q} \tag{2.511}$$

endowed with an automorphism d such that $(d-1)$ is strictly lower triangular

$$(d-1)V^{p,q} \subset \oplus_{r<p,s<q} V^{r,s} \tag{2.512}$$

and a complex conjugation σ which interchanges $V^{p,q}$ and $V^{q,p}$ and fulfills $\sigma d \sigma^{-1} = d^{-1}$.

As explained in Chapter 1 §8.4, the Galois group of the Tannakian category $\mathbb{R}$-MHS of real mixed Hodge structures is the real part of the group

$$G_{MHS} = \mathbb{U}_{MHS} \rtimes (\mathbb{G}_m \times \mathbb{G}_m), \tag{2.513}$$

with respect to the complex conjugation that maps $(u,v) \in \mathbb{G}_m \times \mathbb{G}_m$ to $(\bar{v}, \bar{u})$ and the generator $e_{-n,-m}$ to $-e_{-m,-n}$ where the $e_{n,m}$ are the homogeneous components of $\log d$.

All the computations of this section apply if, in the real case, one replaces the cohomology $H^m(X_v(\mathbb{C}))$ of (2.459) by a real mixed Hodge structure. One gives meaning to the representations (2.454) and (2.455) using the bigrading (2.511) and the complex conjugation σ. The unipotent automorphism d being triangular (2.511), it does not contribute to the trace computation.

We explained in Chapter 1 §8.4 the analogy between the Galois group of the Tannakian category $\mathbb{R}$-MHS of real mixed Hodge structures and the cosmic Galois group that underlies renormalization and classifies flat equisingular connections on the space B.

We saw, in Chapter 1, in the proof of Theorem 1.67, that the flat invariant connections extend automatically to the "compactification" of the base B obtained by the operation (1.339)

$$B \to \underline{B} = B \times_{\mathbb{G}_m} \mathbb{G}_a$$

Moreover the $\mathbb{G}_m$-action on $\underline{B}$ admits a whole set of fixed points, given by (1.340),

$$B_0 = B \times_{\mathbb{G}_m} \{0\} \subset \underline{B} = B \times_{\mathbb{G}_m} \mathbb{G}_a\,.$$

This shows that the base $\underline{B}$ should play a role similar to the adèle class space in the semi-local trace formula of §7.2. It remains to be seen if the corresponding Lefschetz formula is relevant in the framework of renormalization of Chapter 1.

CHAPTER 3

Quantum statistical mechanics and Galois symmetries

1. Overview: three systems

We have seen in the previous chapter how the adèles $\mathbb{A}_{\mathbb{Q}}$ and the noncommutative adèle class space $\mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$ provide a natural geometric setting for a spectral realization of the zeros of the Riemann zeta function and an interpretation as a (semi-local) Lefschetz trace formula of the Weil explicit formula. In this chapter we discuss more in detail the geometry of the adèle class space, in terms of a simple geometric notion: the commensurability relation on $\mathbb{Q}$-lattices.

This formulation leads us to consider more general types of noncommutative adelic quotients and their relation to Galois theory. We follow [**30**], [**86**], [**88**], [**90**], [**91**]. All of the cases discussed in this section are quantum statistical mechanical systems, with nontrivial phase transition phenomena, and with thermodynamical equilibrium states that, at sufficiently low-temperature, recover the points of a classical algebro-geometric moduli space.

The first system we analyze is the Bost–Connes (BC) system [**30**], which is closely related to the adèle class space and is described geometrically in terms of 1-dimensional $\mathbb{Q}$-lattices. As illustrated in the table below, its partition function is the Riemann zeta function, the extremal equilibrium states (KMS states) at sufficiently low-temperature are parameterized by the points of a very simple classical moduli space, the zero-dimensional Shimura variety $Sh(\mathrm{GL}_1, \{\pm 1\})$. The symmetries of the system are given by the group $\mathrm{GL}_1(\hat{\mathbb{Z}})$. The zero-temperature KMS states evaluated on a natural arithmetic subalgebra of the algebra of observables of the system take values that are algebraic numbers and generate the maximal abelian extension $\mathbb{Q}^{\mathrm{cycl}}$ of $\mathbb{Q}$. The class field theory isomorphism intertwines the action of the symmetries and the Galois action on the values of states, thus providing a quantum statistical mechanical reinterpretation of the explicit class field theory of $\mathbb{Q}$.

The second system we present in this chapter is a generalization of the BC system, where instead of considering 1-dimensional $\mathbb{Q}$-lattices, one works with 2-dimensional $\mathbb{Q}$-lattices and their commensurability classes. The corresponding quantum statistical mechanical system was introduced and studied in [**86**] (cf. also [**88**] for a brief summary of the main results of [**86**]). Passing from 1-dimensional to 2-dimensional $\mathbb{Q}$-lattices corresponds, at the level of the corresponding classical algebro-geometric objects, to passing from the Shimura variety of GL_1 to that of GL_2. However, at the quantum statistical mechanical level and in terms of the noncommutative spaces, the GL_2-system exhibits many new properties that were not present in the case of the BC system. One such property is the fact that the symmetries of the system now involve not only automorphisms, but also endomorphisms. Symmetries given by endomorphisms are of crucial importance in order to relate again the quantum statistical mechanical system and its low-temperature KMS states to a rich Galois theory. In this case, this will be the Galois theory of the modular field, which we recall in this chapter. The partition function of the GL_2 system is again related to the Riemann zeta function $\zeta(s)$, this time in the form of a product $Z(\beta) = \zeta(\beta)\zeta(\beta - 1)$. The extremal low-temperature KMS states are parameterized by the points of the classical Shimura variety $Sh(\mathrm{GL}_2, \mathbb{H}^{\pm})$. There is an arithmetic algebra, which is here no longer a subalgebra but an algebra of unbounded multipliers of the algebra of observables, which naturally involves modular functions. For a generic choice of an extremal zero-temperature KMS state, the evaluation on arithmetic elements (multipliers) of the algebra intertwines symmetries of the system, given by the adelic group $\mathbb{Q}^*\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$, and the Galois action of the Galois group of the modular field, realized (via the choice of the state) as an embedded subfield of $\mathbb{C}$.

In both the GL_1- and the GL_2-system, the arithmetic algebra plays a fundamental role. The arithmetic algebras of these two systems can be seen from the point of view of Weil's analogy between trigonometric and elliptic functions developed in [**297**]. The generators of the arithmetic algebra of the BC system can be built out of functions of 1-dimensional lattices and the arithmetic algebra of the GL_2-system contains similar elements based on Eisenstein series and is defined by abstracting the basic properties of these elements. It naturally involves modular functions and Hecke correspondences.

Finally, we discuss another quantum statistical mechanical system, introduced in our joint work with N. Ramachandran [**90**], [**91**]. This system exhibits properties in between the BC and the GL_2 system. This is closely related to the adèle class space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ for $\mathbb{K}$ an imaginary quadratic extension of $\mathbb{Q}$. As the original BC system provides a reformulation of the explicit class field theory of $\mathbb{Q}$, this system will serve the same purpose with respect to the explicit class field theory of $\mathbb{K}$. As is well known, the latter involves in a fundamental way the arithmetic of elliptic curves with complex multiplication. This appears, in our quantum statistical mechanical setting, through the relation to the GL_2-system. Namely, we will see that the system for complex multiplication, based on commensurability of 1-dimensional $\mathbb{K}$-lattices, appears as a subsystem of the GL_2-system, by identifying 1-dimensional $\mathbb{K}$-lattices with a special class of 2-dimensional $\mathbb{Q}$-lattices through the choice of a basis $\{1, \tau\}$ for $\mathbb{K}$ as a vector space over $\mathbb{Q}$. The partition function of this system is the Dedekind zeta function of $\mathbb{K}$. The arithmetic subalgebra is then inherited from the GL_2-system and the values of zero-temperature KMS states and the Galois action are exactly those of class field theory. In this system, as in the GL_2 case, the symmetries are not only given by automorphisms. The presence of symmetries that are given by endomorphisms corresponds to the fact that the field $\mathbb{K}$ may have non-trivial class number.

We also discuss, towards the end of this chapter, some further generalizations of these systems to the case of Shimura varieties [**159**] and to function fields [**172**], [**104**], and the relation of the GL_2-system to the modular Hecke algebras of Connes and Moscovici [**93**], [**94**], [**95**].

The comparative properties of the three systems are illustrated in the table below and will be explained in detail in the rest of this chapter.

System	GL_1	GL_2	$\mathbb{K} = \mathbb{Q}(\sqrt{-d})$
Partition function	$\zeta(\beta)$	$\zeta(\beta)\zeta(\beta-1)$	$\zeta_{\mathbb{K}}(\beta)$
Symmetries	$\mathbb{A}^*_{\mathbb{Q},f}/\mathbb{Q}^*$	$GL_2(\mathbb{A}_{\mathbb{Q},f})/\mathbb{Q}^*$	$\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$
Symmetry group	Compact	Locally compact	Compact
Automorphisms	$\hat{\mathbb{Z}}^*$	$GL_2(\hat{\mathbb{Z}})$	$\hat{\mathcal{O}}^*/\mathcal{O}^*$
Endomorphisms		$GL_2^+(\mathbb{Q})$	$Cl(\mathcal{O})$
Galois group	$Gal(\mathbb{Q}^{ab}/\mathbb{Q})$	$Aut(F)$	$Gal(\mathbb{K}^{ab}/\mathbb{K})$
Extremal KMS_∞	$Sh(GL_1, \pm 1)$	$Sh(GL_2, \mathbb{H}^\pm)$	$\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$

More precisely, in §2 we give a brief introduction to quantum statistical mechanics. We recall the notion of KMS states and its origin motivated by the quantum mechanical analog of the Gibbs measure. We review in §2.1 the general notions of observables, time evolution and Hamiltonian of a quantum statistical mechanical system. We also recall the notion of unbounded multipliers on a C^*-algebra, which will be useful in §§5, 6, 7. We introduce states in §2.2 and the KMS condition, including the case of zero temperature where we discuss the different notions of ground states and KMS_∞ states. We show that extremal KMS states at nonzero temperature form a Choquet simplex. We discuss the extension of states to the multiplier algebra. In §2.3 we discuss symmetries of quantum statistical mechanical systems, induced symmetries on states, both given by automorphisms and by endomorphisms, and the phenomenon of spontaneous symmetry breaking. The induced action of symmetries on zero-temperature KMS states is described in §2.4. The pushforward of KMS states is described in §2.5.
In §3 we introduce the geometric notion of $\mathbb{Q}$-lattices and the relation of commensurability. These are the fundamental notions used in the rest of the chapter to construct the relevant quantum statistical mechanical systems.
We begin in §4 to discuss the geometry of the space of 1-dimensional $\mathbb{Q}$-lattices modulo commensurability. We introduce the groupoid $\mathcal{G}_1$ of the equivalence relation and the one that describes commensurability on $\mathbb{Q}$-lattices up to scaling by $\mathbb{R}^*_+$. We introduce the corresponding C^*-algebras and recall the explicit presentation of the BC algebra $\mathcal{A}_1 = C^*(\mathcal{G}_1/\mathbb{R}^*_+)$. We describe in §4.1 the time evolution induced by the ratio of covolumes of commensurable lattices. In §4.2 we recall the description of the BC systemin terms of Hecke algebras originally given in [**30**]. In §4.3 we show that $\hat{\mathbb{Z}}^*$ acts as symmetries of the BC system. In §4.4 we introduce the arithmetic subalgebra $\mathcal{A}_{1,\mathbb{Q}}$ of $\mathcal{A}_1$, by describing the generators in terms of functions of lattices, their interpretation in terms of trigonometric functions and the relation via Cayley transform to the exponential generators of the BC algebra. We derive explicit division relations between the generators. In §4.5 we give a quick reminder of the Kronecker–Weber theorem and the explicit class field theory of $\mathbb{Q}$. We then

formulate the problem of the quantum statistical mechanical approach to explicit class field theory in §4.6, where we also recall the main result of [**30**] on the classification of KMS states for the BC system and its relation to the Galois theory of the cyclotomic field $\mathbb{Q}^{\rm cycl}$. We discuss in §4.7 the reason why it is not an unreasonable expectation that the approach via algebras and states may prove useful in the explicit class field theory problem. We show in particular how strong is the intertwining property of KMS states between symmetries and Galois action on values of states, by showing how the Galois automorphisms are generally badly behaved from the topological standpoint. We give in §4.8 a reinterpretation of the geometry of the BC system in terms of the Shimura variety of GL_1. We discuss briefly in §4.9 the relation between the algebra of the BC system and the algebra of the "dual system" describing commensurability classes of $\mathbb{Q}$-lattices not up to scaling. This relation will be fully developed in Chapter 4.
We move on to discuss the case of 2-dimensional $\mathbb{Q}$-lattices starting in §5. We describe the reformulation in terms of Tate modules of elliptic curves in §5.1. In §5.2 we introduce the relevant groupoids. Unlike in the 1-dimensional case, here one finds that dividing by the scaling action of $\mathbb{C}^*$ does not preserve the groupoid structure, but one can still define a corresponding convolution algebra $\mathcal{A}_2$. The time evolution induced by the ratio of covolumes is introduced in §5.3, where we also introduce the regular representation of $\mathcal{A}_2$, the associated von Neumann algebra, and the linear functional that gives the KMS_2 state. We discuss the symmetries of the system in §5.4, where we show that the group $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ acts by symmetry, with $\mathrm{GL}_2(\hat{\mathbb{Z}})$ acting by automorphisms and $\mathrm{GL}_2^+(\mathbb{Q})$ by endomorphisms. These combine to give an induced action of $\mathbb{Q}^*\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ on KMS states. We give the explicit form of the action by endomorphisms and of the induced action on KMS states.
In §6 we recall many known facts about the field of modular functions which we need to use in discussing the arithmetic properties of the GL_2-system. We recall explicitly the cases of level one and then more generally level N in §§6.1, 6.2, in particular the Weierstrass $\wp$-function, the j-function, the fact that the field $F_N(\mathbb{C})$ is a finite Galois extension of $\mathbb{C}(j)$, the explicit generators given by the Fricke functions, Eisenstein series, θ-functions, the role of torsion points of elliptic curves and the extensions $F_N(j)$ over $\mathbb{Q}(j)$. General facts about modular functions and modular forms are discussed in §6.3 and some explicit computations for the cases $N = 2$ and $N = 4$ are given in §6.4. Starting with §6.5 we relate the modular field to the geometry of 2-dimensional $\mathbb{Q}$-lattices. We show that an invertible $\mathbb{Q}$-lattice with transcendental j-invariant determines an embedding of the modular field in $\mathbb{C}$ and that all embeddings arise in this way and the same embeddings occur for $\mathbb{Q}$-lattices in the same orbit of the right $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ action. We also discuss the relation between $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ and the automorphisms of the modular field, and the cyclotomic action on the coefficients of the q-expansion.
Section 7 deals with the arithmetic properties of the GL_2-system. We begin in §7.1 by describing some explicit elements of what will be the arithmetic algebra of the GL_2-system. These elements are obtained from Eisenstein series and Hecke correspondences. We then abstract the general properties that we want to require and use them to define the arithmetic algebra $\mathcal{A}_{2,\mathbb{Q}}$ in §7.2. We show that the resulting algebra is a subalgebra of the unbounded multipliers of the C^*-algebra $\mathcal{A}_2$ and is globally invariant under the symmetries of the system described in §5.4. In §7.3 we describe explicit division relations satisfied by the special elements of the arithmetic

subalgebra introduced in §7.1, which generalize the division relations for elliptic functions. A consequence of these relations is the fact that the subalgebra of $\mathcal{A}_{2,\mathbb{Q}}$ generated by these elements is finite-dimensional and reduced over $\mathbb{Q}(j)$; hence it defines an endomotive in the sense discussed in Chapter 4. We begin the classification of KMS states for the GL_2-system in §7.4, where we give a characterization of KMS states as measures on the space of 2-dimensional $\mathbb{Q}$-lattices up to scaling. We show that at low-temperature $\beta > 2$ these measures are supported on the commensurability classes of the invertible $\mathbb{Q}$-lattices. This gives the classification for low-temperature. We compute the partition function of the system and give an explicit formula for the extremal low-temperature KMS states. In §7.5 we describe explicitly the action of symmetries on low-temperature KMS states and in §7.6 we show that the zero-temperature KMS states associated to invertible $\mathbb{Q}$-lattices with transcendental j-invariant intertwine the action of symmetries of the system and the action of the automorphism group of the modular field on the values of states on arithmetic elements. The classification of KMS states then continues in §7.7, by first showing that there are no KMS states in the range $\beta < 1$ and then by the result of Laca–Larsen–Neshveyev on the uniqueness of the KMS states in the intermediate range $1 < \beta < 2$ and the existence of KMS states at $\beta = 1$. As in the case of the BC system, we give in §7.8 a reinterpretation of the geometry of this system in terms of the Shimura variety of GL_2 and we discuss in §7.9 the relation to the noncommutative boundary of modular curves. The compatibility between the BC system and the GL_2-system is described in §7.10.

The final part of the chapter starts in §8 and is dedicated to our joint work with Ramachandran on the quantum statistical mechanics of 1-dimensional $\mathbb{K}$-lattices, for $\mathbb{K}$ an imaginary quadratic field. The main geometric objects, 1-dimensional $\mathbb{K}$-lattices, and the commensurability relation are introduced in §8.1. The relation to 2-dimensional $\mathbb{Q}$-lattices is explained in §8.2. In §8.3 these notions are rewritten is a suitable adelic form, which provides an explicit description of the space of commensurability classes of 1-dimensional $\mathbb{K}$-lattices up to scaling as an adelic quotient. In §8.4 we introduce the corresponding groupoid and C^*-algebra and the time evolution induced by restriction from the GL_2-system. Describing 1-dimensional $\mathbb{K}$-lattices in terms of ideals one finds in §8.5 that the partition function is the Dedekind zeta function . The arithmetic subalgebra is obtained in §8.6 by restriction from the GL_2-system and the symmetries are described explicitly in §8.7 and give the correct group $\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ acting on KMS states. The classification of KMS states of the system is obtained in §8.8, where it is shown that the zero-temperature states have the intertwining property between the symmetries and the Galois action of the Galois group of the maximal abelian extension of $\mathbb{K}$. The low-temperature extremal KMS states, for $\beta > 1$, are parameterized by invertible 1-dimensional $\mathbb{K}$-lattices. For high temperature $\beta \leq 1$ there is a unique KMS state. This is shown in §8.9. Sections 8.9 and 8.10 compare the system for $\mathbb{K}$ with other known systems and quickly describe generalizations of these systems to arbitrary Shimura varieties constructed by Ha and Paugam [**159**].

2. Quantum statistical mechanics

Consider a classical system with Hamiltonian $H(q, p)$ and phase space a symplectic manifold X with local coordinates (q, p) in which the symplectic form ω has the standard Darboux form. In classical statistical mechanics a state for such a Hamiltonian system consists of a probability measure μ on the phase space X, which assigns to each observable f an expectation value, in the form of an average

$$\int_X f\, d\mu. \tag{3.1}$$

In particular, the Hamiltonian $H(q,p)$ and the symplectic structure on the phase space X determine a state, called the *Gibbs canonical ensemble*. It is given by

$$d\mu_G = \frac{1}{Z} e^{-\beta H} d\mu_{\mathrm{Liouville}}, \tag{3.2}$$

normalized by $Z = \int e^{-\beta H} d\mu_{\mathrm{Liouville}}$, where $d\mu_{\mathrm{Liouville}}$ is the Liouville measure associated to the symplectic form ω on X. The Gibbs measure (3.2) depends on a thermodynamic parameter β, which is an inverse temperature, i.e. $\beta = 1/kT$ with k the Boltzmann constant. Here we will set the Boltzmann constant equal to one, so that $\beta = 1/T$.

When passing to infinitely many degrees of freedom, where the interesting phenomena of phase transitions and symmetry breaking occur, the definition of the Gibbs states becomes more involved (cf. [**258**]).

In the quantum mechanical framework, the best replacement for the Gibbs condition is given by the KMS condition at inverse temperature β (cf. [**161**]). We discuss the KMS condition in detail in §2.2 below. It has the advantage that it is simpler in formulation than its classical counterpart given by the Gibbs measure, since it relies only on the involutive algebra $\mathcal{A}$ of observables and its time evolution $\sigma_t \in \mathrm{Aut}(\mathcal{A})$, and does not involve any additional structure, like the symplectic structure on the phase space, nor it requires any delicate argument of approximation by regions of finite volume.

Before giving the general definition in §2.2, we can illustrate the notion of KMS states in the simplest case of finitely many degrees of freedom. Suppose given a finite-dimensional Hilbert space $\mathcal{H} \cong \mathbb{C}^N$. We consider as algebra of observables the matrix algebra $\mathcal{A} = M_n(\mathbb{C})$. Then a state on $\mathcal{A}$ is of the form

$$\varphi(a) = \frac{\mathrm{Tr}(\rho\, a)}{\mathrm{Tr}(\rho)}, \tag{3.3}$$

for some nonzero density matrix $\rho \in M_n(\mathbb{C})$ with $\rho > 0$ (i.e. $\rho = \eta^* \eta$ for some nontrivial $\eta \in M_n(\mathbb{C})$). On the algebra $M_n(\mathbb{C})$ we consider a time evolution generated by a single self-adjoint matrix H (the quantum mechanical Hamiltonian). Namely, we let

$$\sigma_t(a) = e^{itH} a e^{-itH}, \quad \forall t \in \mathbb{R}.$$

Then consider states of the form (3.3) for $\rho = e^{-\beta H}$

$$\varphi_\beta(a) = \frac{\mathrm{Tr}(a\, e^{-\beta H})}{\mathrm{Tr}(e^{-\beta H})} \tag{3.4}$$

for some $\beta \in \mathbb{R}_+^*$. They are equilibrium states, in the sense that they satisfy $\varphi(\sigma_t(a)) = \varphi(a)$ for all $t \in \mathbb{R}$ and for all $a \in \mathcal{A}$. Moreover, such Gibbs states are solutions to a variational problem with respect to the entropy functional (cf. [**160**]). The notion of KMS states, which we discuss in detail in the rest of this section, generalizes this simple example to the infinite-dimensional case.

2.1. Observables and time evolution.

While in the classical setting an observable is just a function on the phase space X of the Hamiltonian system, in the quantum mechanical setting the observables form a (generally noncommutative) C^*-algebra $\mathcal{A}$. Moreover the Hamiltonian is replaced by its associated Hamiltonian flow, namely the continuous group homomorphism

$$\sigma : \mathbb{R} \to \mathrm{Aut}(\mathcal{A}), \tag{3.5}$$

which specifies the time evolution.

Thus, in the context of quantum statistical mechanics, instead of assigning data (X, H) of a phase space and a Hamiltonian as in the classical case, one specifies as data a C^*-dynamical system $(\mathcal{A}, \sigma)$ given by the algebra of observables (a separable C^*-algebra $\mathcal{A}$) and the time evolution (3.5). Thus, in the following we often refer to the data $(\mathcal{A}, \sigma)$ as a quantum statistical mechanical system.

One can represent such a quantum statistical mechanical system concretely in a Hilbert space $\mathcal{H}$ using the following notions

Definition 3.1. *We say that the representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ of $\mathcal{A}$ in the algebra of bounded operators on $\mathcal{H}$ is*

- *A* covariant *representation if there exists a self-adjoint operator H on $\mathcal{H}$ such that*

$$\pi(\sigma_t(a)) = e^{itH}\pi(a)e^{-itH}, \quad \forall t \in \mathbb{R}. \tag{3.6}$$

- *A* positive energy *representation if the Hamiltonian H of* (3.6) *can be chosen with spectrum* $\mathrm{Spec}(H) \subset [0, \infty)$.

In the case when the C^*-algebra $\mathcal{A}$ is non-unital, we consider the "one-point compactification" $\mathcal{A}^{\mathrm{comp}}$ of $\mathcal{A}$, i.e. the C^*-algebra obtained by adding a unit to $\mathcal{A}$,

$$\mathcal{A}^{\mathrm{comp}} = \mathcal{A} \oplus \mathbb{C} \tag{3.7}$$

with product given by $(a, \lambda)(b, \mu) = (ab + \lambda b + \mu a, \lambda\mu)$ for a, $b \in \mathcal{A}$ and $\lambda, \mu \in \mathbb{C}$. In the following we will consider another possible compactification, which corresponds in the commutative case to the Stone–Čech compactification. Recall that an essential ideal $\mathcal{I}$ in a C^*-algebra $\mathcal{A}$ is a closed ideal which has nontrivial intersection with every other nonzero closed ideal of $\mathcal{A}$.

Definition 3.2. *Given a (non-unital) C^*-algebra $\mathcal{A}$ there exists a maximal C^*-algebra $M(\mathcal{A})$ with the property that $\mathcal{A} \subset M(\mathcal{A})$ is an essential ideal. The algebra $M(\mathcal{A})$ is unique up to isomorphism and is called the* multiplier algebra *of $\mathcal{A}$.*

In the commutative case, when $\mathcal{A} = C_0(X)$ for some locally compact Hausdorff space X, one has $M(\mathcal{A}) = C(X^{\mathrm{SC}})$ where X^{SC} is the Stone–Čech compactification of X (cf. e.g. [**291**]).

We can equivalently view a bounded multiplier on a C^*-algebra $\mathcal{A}$ as a bounded linear operator $T : \mathcal{A} \to \mathcal{A}$ with the property that $T(ab) = aT(b)$ for all $a, b \in \mathcal{A}$. Thus, one can also introduce the notion of *unbounded multipliers* of a C^*-algebra $\mathcal{A}$.

Definition 3.3. *An unbounded multiplier on a topological algebra $\mathcal{A}$ is a linear operator $T : \mathcal{D}(T) \to \mathcal{A}$, defined on a dense ideal $\mathcal{D}(T) \subset \mathcal{A}$, with the property that $T(ab) = aT(b)$ for all $a \in \mathcal{D}(T)$ and all $b \in \mathcal{A}$.*

In the commutative case, with $\mathcal{A} = C_0(X)$, an unbounded multiplier on $\mathcal{A}$ is given by multiplication by some (unbounded) continuous function on X (cf. [**304**]).

2.2. The KMS condition.

In the commutative case, states on a commutative C^*-algebra $C(X)$ are the probability measures on the compact topological space X. Thus, one can generalize the notion of a probability measure to noncommutative spaces, by considering the following notion of states on the algebra $\mathcal{A}$ of observables.

DEFINITION 3.4. *Let $\mathcal{A}$ be a unital C^*-algebra. A state φ on $\mathcal{A}$ is a continuous linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ satisfying the normalization and positivity conditions*

$$\varphi(1) = 1, \quad \text{and} \quad \varphi(a^* a) \geq 0. \tag{3.8}$$

When the C^*-algebra $\mathcal{A}$ is non-unital, one replaces the condition $\varphi(1) = 1$ in (3.8) by the normalization condition $\|\varphi\| = 1$ with

$$\|\varphi\| := \sup_{x \in \mathcal{A},\, \|x\| \leq 1} |\varphi(x)| \,. \tag{3.9}$$

In the non-unital case, by Proposition 2.1.5 of [**120**] the states (Definition 3.4) are restrictions of states on the unital C^*-algebra $\mathcal{A}^{\mathrm{comp}}$ obtained by adjoining a unit (the one-point compactification) as in (3.7).

PROPOSITION 3.5. *Given a representation of a unital C^*-algebra $\mathcal{A}$ on a Hilbert space $\mathcal{H}$, namely a $*$-homomorphism $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$, one obtains states on $\mathcal{A}$ by setting*

$$\varphi(a) = \langle \xi, \pi(a)\, \xi \rangle,$$

for $\xi \in \mathcal{H}$ with $\|\xi\| = 1$. Conversely, a state φ on a unital C^*-algebra $\mathcal{A}$ determines a representation of $\mathcal{A}$ on a Hilbert space, via the GNS (Gelfand–Naimark–Segal) construction, where the bilinear form $\langle a, b \rangle = \varphi(a^*\, b)$ defines an inner product on $\mathcal{H} = \mathcal{A}/\mathcal{I}$, for the left ideal $\mathcal{I} = \{a : \varphi(a^* a) = 0\}$. The algebra $\mathcal{A}$ acts on $\mathcal{H}$ by left multiplication and the element $\xi = 1 + \mathcal{I}$ is a cyclic vector for the representation. The state is then written in the form $\varphi(a) = \langle \xi, a\, \xi \rangle$.

We refer to [**120**] Proposition 2.4.4 for the proof.

In the presence of a time evolution σ_t on the C^*-algebra $\mathcal{A}$, there is a class of states that are particularly interesting, as they describe the *equilibrium states* of the system $(\mathcal{A}, \sigma_t)$, depending on a thermodynamical parameter β. This parameter is an inverse temperature $\beta = 1/T$ where one sets the Boltzmann constant equal to one. The equilibrium states are defined by the KMS (Kubo–Martin–Schwinger) condition as follows (cf. [**33**], [**160**], [**161**]).

DEFINITION 3.6. *Suppose given a C^*-dynamical system $(\mathcal{A}, \sigma_t)$, that is, a C^*-algebra $\mathcal{A}$ together with a 1-parameter group of automorphisms $\sigma : \mathbb{R} \to \mathrm{Aut}(\mathcal{A})$. For a given $0 < \beta < \infty$, a state φ on the unital C^*-algebra $\mathcal{A}$ satisfies the KMS condition at inverse temperature β if for all $a, b \in \mathcal{A}$, there exists a function $F_{a,b}(z)$ which is holomorphic on the strip*

$$I_\beta = \{z \in \mathbb{C} \,|\, 0 < \Im(z) < \beta\}, \tag{3.10}$$

continuous on the boundary ∂I_β and bounded, with the property that for all $t \in \mathbb{R}$

$$F_{a,b}(t) = \varphi(a \sigma_t(b)) \quad \text{and} \quad F_{a,b}(t + i\beta) = \varphi(\sigma_t(b) a). \tag{3.11}$$

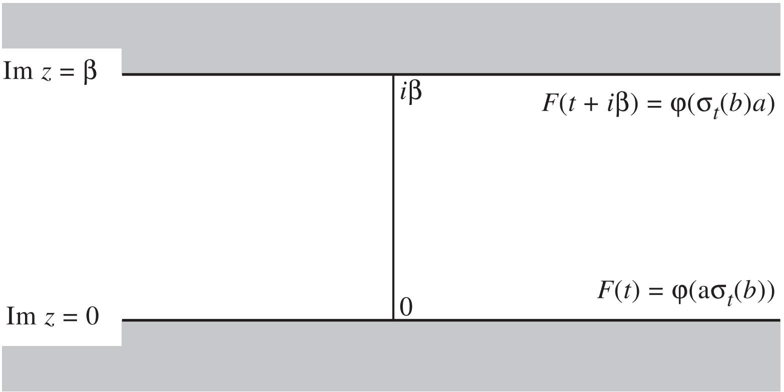


FIGURE 1. The strip I_β in the KMS condition

These conditions imply the invariance $\varphi(\sigma_t(a)) = \varphi(a)$ of the state φ. This can be seen as a consequence of the Liouville Theorem which asserts that the only bounded entire functions are the constants.
The set Σ_β of KMS states at $\beta < \infty$ is a compact convex set for the weak topology:

$$\varphi_n \to \varphi \Leftrightarrow \varphi_n(a) \to \varphi(a) \quad \forall a \in \mathcal{A}$$

Thus, it makes sense to consider the set $\mathcal{E}_\beta$ of its extremal points.

One can also define equilibrium states at zero-temperature (i.e. for $\beta = \infty$). There are two different notions in this case. Unlike other references in the literature, we reserve the terminology KMS$_\infty$ states only for the stronger condition, while we refer to the weaker notion more generally as "ground states".

DEFINITION 3.7. *Suppose given a C^*-dynamical system $(\mathcal{A}, \sigma_t)$. A state φ on $\mathcal{A}$ is a* KMS$_\infty$ state *if it is a weak limit of KMS$_\beta$ states as $\beta \to \infty$. Namely, there exist KMS$_\beta$ states φ_β, for sufficiently large β, such that for all $a \in \mathcal{A}$*

$$\varphi(a) = \lim_{\beta\to\infty} \varphi_\beta(a). \tag{3.12}$$

A state φ on the unital C^*-algebra $\mathcal{A}$ is a ground state *if for all $a, b \in \mathcal{A}$ there exists a bounded holomorphic function $F_{a,b}(z)$ on the upper half-plane, such that $F_{a,b}(t) = \varphi(a\sigma_t(b))$.*

When φ is a ground state it is invariant (one can use the Schwarz reflection principle applied to $F_{1,b}(t) = \varphi(\sigma_t(b))$ with $b = b^*$). The corresponding GNS representation π_φ (cf. Proposition 2.4.4 of [**120**]) is covariant and of positive energy according to Definition 3.1.
It is clear that the notion of ground states is much weaker than that of KMS$_\infty$ states adopted in Definition 3.7. In fact, in the simplest case of the trivial time evolution

one can see that *all states* are ground states according to Definition 3.7, while only *tracial states* satisfy (3.11) at $\beta < \infty$.
One reason for the above refinement of the notion of ground states at zero-temperature is to preserve the simplicial structure of the set of KMS_β states which is ensured at non-zero-temperature by:

PROPOSITION 3.8. *Let $(\mathcal{A}, \sigma)$ be a C^*-dynamical system with a unital C^*-algebra $\mathcal{A}$.*

- *A KMS_β state for $0 < \beta < \infty$ is extremal if and only if the corresponding GNS representation is factorial.*
- *For any $0 < \beta < \infty$, the set Σ_β of KMS_β states is a convex compact Choquet simplex .*
- *The set Σ_∞ of KMS_∞ states defined as in Definition 3.7 is also a convex compact set.*

PROOF. We only give a brief sketch of the argument and refer to [**33**] for more details. First let φ be a KMS_β state and assume that the corresponding GNS representation π_φ is factorial. Let $M = \pi_\varphi(\mathcal{A})''$ be the corresponding factor. If φ fails to be extremal there exists a positive linear form $\psi < \varphi$ on $\mathcal{A}$ which is KMS_β and is not a scalar multiple of φ. Then ψ extends to M by weak continuity and the corresponding functional $\tilde{\psi}$ on M fulfills the same KMS_β condition as the extension $\tilde{\varphi}$ of φ to M. It follows that $\tilde{\psi}(x) = \tilde{\varphi}(hx)$ for some positive element h in the center of M, but since M is a factor it follows that ψ is in fact proportional to φ.
Conversely if $M = \pi_\varphi(\mathcal{A})''$ fails to be a factor one obtains a non-trivial decomposition of φ using functionals of the form $\psi(x) = \varphi(hx)$ for some positive element h in the center of M.

The second statement follows since the decomposition of a KMS_β state into extremal KMS_β states coincides with the decomposition of the state into factor states, and the latter is unique.

The weak limit of convex compact sets is convex and compact. There is however no guarantee in general that a weak limit of simplices be a simplex. This however is obviously the case when the limit is stationary (i.e. $\Sigma_\infty = \Sigma_\beta$ for $\beta \geq \beta_0$) which will be the case in our examples.

∎

We refer to the books [**33**] and [**160**] for a more detailed discussion of KMS states and phase transitions.

The typical pattern for a system with a single phase transition is that the simplex Σ_β consists of a single point for $\beta \leq \beta_c$, i.e. when the temperature is larger than the critical temperature T_c, and is non-trivial (of some higher dimension in general) when the temperature lowers. Systems can exhibit a more complex pattern of multiple phase transitions, where no KMS state exists above a certain temperature. The GL_2 system, which we analyze in detail in this chapter, is an example of a system that exhibits a more elaborate behavior.

In the nonunital case, the KMS condition is defined in the same way, using the condition (3.11).

DEFINITION 3.9. *Given a C^*-dynamical system $(\mathcal{A}, \sigma)$, let $M(A)_\sigma \subset M(A)$ be the C^*-subalgebra of elements in the multiplier algebra $M(A)$ such that $t \mapsto \sigma_t(x)$ is norm continuous.*

We then have the following result.

PROPOSITION 3.10. *Any state φ on $\mathcal{A}$ admits a canonical extension to a state, still denoted by φ, on the multiplier algebra $M(\mathcal{A})$ of $\mathcal{A}$. The canonical extension to $M(\mathcal{A})$ of a KMS state still satisfies the KMS condition on the subalgebra $M(\mathcal{A})_\sigma$.*

PROOF. For the first statement we refer to [**241**]. The proof of the second statement illustrates a general density argument, where the continuity of $t \mapsto \sigma_t(x)$ is used to control the uniform continuity in the closed strip I_β, in order to apply the Montel theorem of normal families. Indeed, by weak density of $\mathcal{A}$ in $M(\mathcal{A})$, one obtains a sequence of holomorphic functions, but one only controls their uniform continuity on smooth elements of $M(\mathcal{A})_\sigma$.

□

As will become clear in the following sections, it is useful to extend whenever possible the integration provided by a state to unbounded multipliers of A. In the commutative case, this corresponds to the fact that states are indeed given by integration with respect to a probability measure and one usually finds that the class of integrable functions not only contains bounded functions on the space X but also a certain class of unbounded integrable functions.

We have seen in Proposition 3.8 that in the unital case the set of KMS states at a fixed temperature is a convex compact Choquet simplex (possibly empty and in general infinite dimensional). In the nonunital case, given a σ_t-invariant subalgebra $\mathcal{C} \subset M(\mathcal{A})_\sigma$, $\mathcal{C} \supset \mathcal{A}$, the set $\Sigma_\beta(\mathcal{C})$ of KMS_β states on $\mathcal{C}$ should be viewed as a compactification of the set of KMS_β states on $\mathcal{A}$. The restriction from $\mathcal{C}$ to $\mathcal{A}$ maps $\Sigma_\beta(\mathcal{C})$ to KMS_β positive linear forms on $\mathcal{A}$ of norm less than or equal to one. These are referred to as *quasi-states*.

2.3. Symmetries.

We now discuss the notion of symmetries of a quantum statistical mechanical system, which will play a crucial role in the rest of this chapter.

DEFINITION 3.11. *A locally compact group G is a group of* symmetries by automorphisms *of a quantum statistical mechanical system $(\mathcal{A}, \sigma)$ if there is a faithful representation*

$$\alpha : G \to \mathrm{Aut}(\mathcal{A}, \sigma_t), \tag{3.13}$$

that is, a continuous embedding $\alpha : G \hookrightarrow \mathrm{Aut}(\mathcal{A})$ such that

$$\sigma_t \alpha_g = \alpha_g \sigma_t, \quad \forall\, g \in G\,, \forall t \in \mathbb{R}\,. \tag{3.14}$$

A group $G \subset \mathrm{Aut}(\mathcal{A}, \sigma_t)$ of symmetries by automorphisms acts on KMS states by pull back:

$$g^*\varphi(a) = \varphi(g(a)). \tag{3.15}$$

and one has by construction

$$(g_1 g_2)^*\varphi = g_2^*(g_1^*\varphi) \;\; \forall g_j \in G\,. \tag{3.16}$$

We denote by $\mathrm{Inn}(\mathcal{A}, \sigma_t)$ the normal subgroup of $\mathrm{Aut}(\mathcal{A}, \sigma_t)$ of inner automorphisms of $\mathcal{A}$ that commute with the time evolution.

Let $\mathcal{U}^\sigma(\mathcal{A})$ denote the unitary group of the fixed point algebra of σ_t, that is,

$$\mathcal{U}^\sigma(\mathcal{A}) = \{u \in \mathcal{A} \,:\, u^*u = uu^* = 1,\ \sigma_t(u) = u, \forall t \in \mathbb{R}\}. \tag{3.17}$$

LEMMA 3.12. *Let $G \subset \mathrm{Aut}(\mathcal{A})$ be a group of symmetries by automorphisms of the system $(\mathcal{A}, \sigma)$. Let $G_0 \lhd G$ be a normal subgroup such that $G_0 \subset \mathcal{U}^\sigma(\mathcal{A})$. Then there is an induced action of the quotient group G/G_0 on KMS states.*

PROOF. The group $\mathcal{U}^\sigma(\mathcal{A})$ acts by inner automorphisms of the dynamical system $(\mathcal{A}, \sigma)$. Namely, an element $u \in \mathcal{U}^\sigma(\mathcal{A})$ acts as

$$\mathrm{ad}(u)\, x := uxu^*, \qquad \forall x \in \mathcal{A}\,. \tag{3.18}$$

Given a C^*-dynamical system $(\mathcal{A}, \sigma)$, a KMS_β state φ has the property that

$$\varphi(b\, \sigma_{i\beta}(a)) = \varphi(ab), \tag{3.19}$$

for all a, b in a norm dense and σ-invariant $*$-subalgebra of $\mathcal{A}$. In fact, condition (3.19) is equivalent to the KMS condition of (3.11) (cf. Definition 5.3.1 and Corollary 5.3.7 of [**33**]).
In particular, this shows that the elements of $\mathcal{U}^\sigma(\mathcal{A})$ act trivially on KMS states since

$$\varphi(uxu^*) = \varphi(xu^*\sigma_{i\beta}(u)) = \varphi(x), \qquad \forall x, \tag{3.20}$$

where $\sigma_{i\beta}(a)$ denotes the analytic continuation $z \mapsto \sigma_z(a)$ of $t \mapsto \sigma_t(a)$ for for a in the dense subalgebra of $\mathcal{A}$. This completes the proof. $\square$

It is important to consider a more general type of symmetries, besides those given by automorphisms. In fact, as we are going to see in the cases of the GL(2) system and of the system for imaginary quadratic fields, it is crucial to take into account also symmetries by *endomorphisms*.
This type of symmetry plays a key role in the theory of superselection sectors in algebraic quantum field theory developed by Doplicher–Haag–Roberts (cf.[**160**], Chapter IV).

DEFINITION 3.13. *An* endomorphism *ρ of a C^*-dynamical system $(\mathcal{A}, \sigma_t)$ is a $*$-homomorphism $\rho : \mathcal{A} \to \mathcal{A}$ that commutes with σ_t*

$$\sigma_t\, \rho = \rho\, \sigma_t, \qquad \forall t \in \mathbb{R}\,. \tag{3.21}$$

It follows then that $\rho(1) = e$ is an idempotent fixed by σ_t,

$$\sigma_t(e) = e, \qquad \forall t \in \mathbb{R}\,. \tag{3.22}$$

A KMS state φ can be pulled back along an endomorphism ρ, provided that the element $e = \rho(1)$ satisfies $\varphi(e) \neq 0$. In that case, one has the pullback of the form

$$\rho^*(\varphi)(a) := \frac{\varphi \circ \rho(a)}{\varphi(e)}\,, \qquad \forall a \in \mathcal{A}. \tag{3.23}$$

Exactly as we considered unitaries fixed by the time evolution, $u \in \mathcal{U}^\sigma(\mathcal{A})$, in the case of automorphisms, we now consider isometries that are eigenvectors of the time evolution, namely we consider the set

$$\mathcal{I}^\sigma(\mathcal{A}) := \{u \in \mathcal{A} \ \text{ such that } \ u^*u = 1, \ \text{ and } \ \sigma_t(u) = \lambda^{it}\, u,\ \forall t \in \mathbb{R}\,\}, \tag{3.24}$$

for some $\lambda \in \mathbb{R}_+^*$ with $\lambda \geq 1$.

An element $u \in \mathcal{I}^\sigma(\mathcal{A})$ defines an *inner endomorphism* $\mathrm{ad}(u)$ of the dynamical system $(\mathcal{A}, \sigma)$, which acts by

$$\mathrm{ad}(u)x := uxu^*, \quad \forall x \in \mathcal{A}\,. \tag{3.25}$$

One obtains then the following result.

LEMMA 3.14. *The inner endomorphisms of the dynamical system $(\mathcal{A}, \sigma_t)$ act trivially on the set of KMS$_\beta$ states, namely*

$$ad(u)^*(\varphi) = \varphi, \quad \forall \varphi \in \Sigma_\beta\,. \tag{3.26}$$

PROOF. We proceed as in the proof of Lemma 3.12 and we obtain that the KMS$_\beta$ condition (3.19) shows that $\varphi(uu^*) = \lambda^{-\beta} > 0$ so that the pullback $\mathrm{ad}(u)^*(\varphi)$ is well defined. Following the same argument of Lemma 3.12, using the KMS$_\beta$ condition (3.19) applied to the pair $b = xu^*$ and $a = u$, we obtain as in (3.20) that $\mathrm{ad}(u)^*(\varphi) = \varphi$. ☐

REMARK 3.15. *The inequality $\lambda \geq 1$ follows from the existence of a KMS$_\beta$ state. Indeed $\varphi(uu^*) = \lambda^{-\beta} > 0$ and since $uu^* \leq 1$ one gets $\lambda^{-\beta} \leq 1$ and $\lambda \geq 1$.*

2.4. Warming up and cooling down.

At temperature $T = 0$ (i.e. $\beta = \infty$) it is in general no longer true that the endomorphisms act directly on the set Σ_∞ of KMS$_\infty$ states, even if the action is well defined for all the $\beta < \infty$. In fact, one can easily find examples where the condition $\varphi(e) \neq 0$ that makes it possible to pull back states along an endomorphism ρ with $e = \rho(1)$ is satisfied (as a condition on $\varphi \in \Sigma_\beta$) for all $\beta < \infty$ but where in the limit one has $\lim_{\beta\to\infty} \varphi_\beta(e) = 0$.

In such cases it is sometime still possible to define a nontrivial action on KMS$_\infty$ states, through a process we refer to as *warming up and cooling down*. For such process to apply, one needs an assumption about the structure of the set Σ_β of extremal KMS states that is often satisfied in significant cases.

Consider the *warming up map* $W_\beta : \Sigma_\infty \to \Sigma_\beta$ defined as

$$W_\beta(\varphi)(x) = \frac{\mathrm{Tr}(\pi(x)\, e^{-\beta H})}{\mathrm{Tr}(e^{-\beta H})}, \quad \forall x \in \mathcal{A}\,, \tag{3.27}$$

where π is the GNS representation associated to the KMS$_\infty$ state φ and H is the positive energy Hamiltonian, implementing the time evolution in the representation π.

We assume that for sufficiently large $\beta \geq \beta_0$ the map $W_\beta : \Sigma_\infty \to \Sigma_\beta$ is well defined and is a homeomorphism.

Under this assumption, if for $\beta_0 \leq \beta < \infty$ one has a well defined action of an endomorphism ρ on Σ_β by pullback (i.e. that $\varphi_\beta(e) \neq 0$), then we can define an action on Σ_∞ by first "warming up" through the map W_β, then taking the action on Σ_β and then "cooling down" by the inverse map.

2.5. Pushforward of KMS states.

We have seen in §2.3 above that KMS states can be pulled back along automorphisms or endomorphisms of the algebra $\mathcal{A}$ compatible with the time evolution. Here we discuss another important functoriality, namely the pushforward of KMS states under equivariant surjections, modulo normalization.

PROPOSITION 3.16. *Let $(\mathcal{A}, \sigma)$ be a C^*-dynamical system and $\mathcal{J}$ a norm closed two-sided ideal of $\mathcal{A}$ globally invariant under σ_t. Let u_n be a quasi-central approximate unit for $\mathcal{J}$. For any KMS$_\beta$ state φ on $(\mathcal{A}, \sigma_t)$ the sequence $\varphi((1-u_n)\,x)$ converges, for all $x \in \mathcal{A}$ and defines a KMS$_\beta$ positive linear form on $(\mathcal{A}/\mathcal{J}, \sigma_t)$,*

$$\psi(x) = \lim_{n\to\infty} \varphi((1-u_n)\,x), \qquad \forall x \in \mathcal{A}\,. \tag{3.28}$$

PROOF. Let $\mathcal{A}''$ be the double dual of $\mathcal{A}$ and $p \in \mathcal{A}''$ the central open projection corresponding to the ideal $\mathcal{J}$ (cf. [**241**]). By construction, the u_n converge weakly to p (cf. [**241**] 3.12.14) so that the convergence follows as well as the positivity of ψ. By construction ψ vanishes on $\mathcal{J}$. To get the KMS$_\beta$ condition one applies (3.11) with $a = (1-u_n)\,x$ and $b = y$, where y is in a dense subalgebra of $\mathcal{A}$. Then one gets a bounded uniformly continuous sequence $F_n(z)$ of holomorphic functions in the strip I_β of (3.10) with

$$F_n(t) = \varphi((1-u_n)\,x\,\sigma_t(y)) \quad \text{and} \quad F_n(t+i\beta) = \varphi(\sigma_t(y)(1-u_n)\,x) \quad \forall t \in \mathbb{R}.$$

Using the Montel theorem on normal families and the quasi-central property of u_n one gets the KMS$_\beta$ condition for ψ. ☐

3. $\mathbb{Q}$-lattices and commensurability

We introduce the main geometric notion that underlies the construction of the quantum statistical mechanical systems of interest in this chapter.

DEFINITION 3.17. *A $\mathbb{Q}$-lattice in $\mathbb{R}^n$ consists of a pair (Λ, ϕ) of a lattice $\Lambda \subset \mathbb{R}^n$ (a cocompact free abelian subgroup of $\mathbb{R}^n$ of rank n) together with a homomorphism of abelian groups*

$$\phi : \mathbb{Q}^n/\mathbb{Z}^n \longrightarrow \mathbb{Q}\Lambda/\Lambda. \tag{3.29}$$

A $\mathbb{Q}$-lattice is invertible *if ϕ is an isomorphism.*
Two $\mathbb{Q}$-lattices are commensurable,

$$(\Lambda_1, \phi_1) \sim (\Lambda_2, \phi_2),$$

if and only if

$$\mathbb{Q}\Lambda_1 = \mathbb{Q}\Lambda_2 \quad \textit{and} \quad \phi_1 = \phi_2 \mod \Lambda_1 + \Lambda_2. \tag{3.30}$$

Let $\mathcal{L}_n$ denote the set of commensurability classes of n-dimensional $\mathbb{Q}$-lattices.

LEMMA 3.18. *The relation of commensurability between n-dimensional $\mathbb{Q}$-lattices is an equivalence relation.*

PROOF. Let $L_j = (\Lambda_j, \phi_j)$, for $j = 1, 2, 3$, be n-dimensional $\mathbb{Q}$-lattices. Assume that the pairs L_1, L_2 and L_2, L_3 are commensurable. This means that the lattices Λ_j are commensurable and are of finite index in

$$\Lambda = \Lambda_1 + \Lambda_2 + \Lambda_3,$$

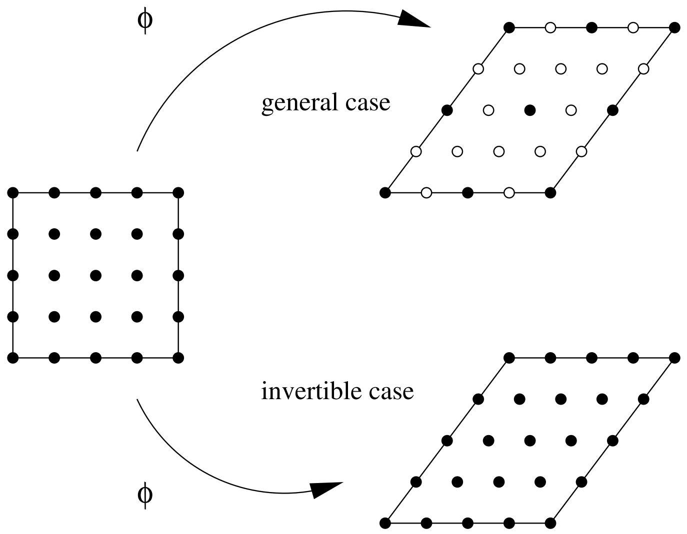


FIGURE 2. Generic and invertible 2-dimensional $\mathbb{Q}$-lattices.

which is indeed a lattice by commensurability. Moreover, one has $\phi_1 - \phi_2 = 0$ mod Λ and $\phi_2 - \phi_3 = 0$ mod Λ, hence also $\phi_1 - \phi_3 = 0$ mod Λ. Notice that $\Lambda' = \Lambda_1 + \Lambda_3$ is of finite index in Λ, hence $\phi_1 - \phi_3$ gives a group homomorphism

$$\mathbb{Q}^n/\mathbb{Z}^n \to \Lambda/\Lambda'.$$

This has to be trivial, since $\mathbb{Q}^n/\mathbb{Z}^n$ is infinitely divisible while Λ/Λ' is finite. This shows that $\phi_1 - \phi_3 = 0$ modulo $\Lambda' = \Lambda_1 + \Lambda_3$, so that we conclude that the pair L_1, L_3 is also commensurable. □

The datum of the homomorphism (3.29) can be thought of as assigning a labeling of the torsion points of the lattice Λ. However, it is important that we do not require such a labeling to be one-to-one. In fact, ϕ is only assumed to be a homomorphism and it need not be an isomorphism. It is precisely because we allow for the possibility of "degenerate" labeling systems that fail to be isomorphisms that the commensurability relation gives rise to a quotient $\mathcal{L}_n$ that has the typical properties of a noncommutative space. This can be seen by the fact that the set $\mathcal{L}_n$ has the cardinality of the continuum but one cannot construct a countable collection of measurable functions that separate points.

REMARK 3.19. *Two invertible $\mathbb{Q}$-lattices are commensurable if and only if they are equal. Most $\mathbb{Q}$-lattices are not commensurable to an invertible one. Consider the set $\mathcal{L}_n^0 \subset \mathcal{L}_n$ of $\mathbb{Q}$-lattices that are commensurable to an invertible $\mathbb{Q}$-lattice. Then the set of invertible $\mathbb{Q}$-lattices gives a cross-section of the equivalence relation on $\mathcal{L}_n^0$.*

Another notion we will need in the following is that of divisibility for $\mathbb{Q}$-lattices.

DEFINITION 3.20. *For an n-dimensional $\mathbb{Q}$-lattice (Λ, ϕ), let ϕ_N denote the restriction of the map ϕ to the N-torsion points of $(\mathbb{Q}/\mathbb{Z})^n$,*

$$\phi_N : \left(\frac{1}{N}\mathbb{Z}/\mathbb{Z}\right)^n \to \frac{1}{N}\Lambda/\Lambda. \tag{3.31}$$

The $\mathbb{Q}$-lattice (Λ, ϕ) is divisible by N if $\phi_N = 0$.

In the following we concentrate on the cases of 1-dimensional and of 2-dimensional $\mathbb{Q}$-lattices. We study the noncommutative geometry of the spaces $\mathcal{L}_1$ and $\mathcal{L}_2$ and of the corresponding spaces $\mathcal{L}_1/\mathbb{R}_+^*$ and $\mathcal{L}_2/\mathbb{C}^*$ of commensurability classes of $\mathbb{Q}$-lattices up to scaling. The latter provide noncommutative generalizations of the Shimura varieties $Sh(\mathrm{GL}_1, \pm)$ and $Sh(\mathrm{GL}_2, \mathbb{H}^{\pm})$, in a sense that will be made precise later in this section. These noncommutative spaces support interesting quantum statistical mechanical systems which are related, respectively, to the explicit class field theory of $\mathbb{Q}$ and to the Galois theory of the field of modular functions (cf. [**30**], [**86**], [**88**], [**91**]).

4. 1-dimensional $\mathbb{Q}$-lattices

We begin by giving an explicit description of the data parameterizing $\mathbb{Q}$-lattices in dimension one.

LEMMA 3.21. *1) The space of 1-dimensional $\mathbb{Q}$-lattices can be identified with the product $\hat{\mathbb{Z}} \times \mathbb{R}_+^*$ by the map*

$$\ell(\rho, \lambda) = (\lambda^{-1}\mathbb{Z}, \lambda^{-1}\rho) \tag{3.32}$$

2) The groupoid $\mathcal{G}_1$ of the equivalence relation of commensurability on 1-dimensional $\mathbb{Q}$-lattices is isomorphic to the étale groupoid

$$\mathcal{G}_1 = \{(r, \rho, \lambda) \,|\, r \in \mathbb{Q}_+^*,\ \lambda \in \mathbb{R}_+^*,\ \rho \in \hat{\mathbb{Z}},\ \text{ such that } r\rho \in \hat{\mathbb{Z}}\}, \tag{3.33}$$

with source and range maps $(r, \rho, \lambda) \mapsto (\rho, \lambda)$ and $(r, \rho, \lambda) \mapsto (r\rho, r\lambda)$, and composition

$$(r_1, \rho_1, \lambda_1) \circ (r_2, \rho_2, \lambda_2) = (r_1 r_2, \rho_2, \lambda_2), \qquad \text{if } r_2\rho_2 = \rho_1\,,\ r_2\lambda_2 = \lambda_1. \tag{3.34}$$

3) The space of 1-dimensional $\mathbb{Q}$-lattices up to scaling is identified with $\hat{\mathbb{Z}}$.

PROOF. 1) A 1-dimensional $\mathbb{Q}$-lattice (Λ, ϕ) can always be described in the following way. A 1-dimensional lattice $\Lambda \subset \mathbb{R}$ is always of the form $\Lambda = \lambda^{-1}\mathbb{Z}$, for some $\lambda \in \mathbb{R}_+^*$. Having described Λ uniquely in this form, the remaining datum $\phi : \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}\Lambda/\Lambda$ is then of the form $\phi = \lambda^{-1}\rho$, with the same $\lambda \in \mathbb{R}_+^*$ and for some $\rho \in \hat{\mathbb{Z}}$. Here we use the identification

$$\hat{\mathbb{Z}} \cong \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, \mathbb{Q}/\mathbb{Z}), \tag{3.35}$$

where the topology on the right hand side is that of pointwise convergence.
2) To $g = (r, \rho, \lambda) \in \mathcal{G}_1$ one associates the pair of $\mathbb{Q}$-lattices corresponding by the map ℓ of (3.32) to the range and source of g, i.e. the pair

$$(\ell(r\rho, r\lambda), \ell(\rho, \lambda)) = ((r^{-1}\lambda^{-1}\mathbb{Z}, \lambda^{-1}\rho), (\lambda^{-1}\mathbb{Z}, \lambda^{-1}\rho))$$

This is a pair of commensurable $\mathbb{Q}$-lattices and any such pair is obtained in this way. The composition in $\mathcal{G}_1$ corresponds to the composition in the groupoid of the commensurability equivalence relation.
3) When considering $\mathbb{Q}$-lattices up to scaling, one mods out by the action of $\mathbb{R}_+^*$. The lattice up to scale is then uniquely identified with $\mathbb{Z}$ and one is left only with the datum of the homomorphism $\rho : \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$, i.e. with assigning an element $\rho \in \hat{\mathbb{Z}}$. $\square$

We now consider the quotient $\mathcal{G}_1/\mathbb{R}_+^*$ by the action of scaling. Since the scaling action of $\mathbb{R}_+^*$ on 1-dimensional $\mathbb{Q}$-lattices is free, the quotient $\mathcal{G}_1/\mathbb{R}_+^*$ is still a groupoid. We let $\mathcal{U}_1$ be the étale groupoid of pairs

$$\mathcal{U}_1 = \{(r, \rho) \,|\, r \in \mathbb{Q}_+^*,\ \rho \in \hat{\mathbb{Z}},\ \text{ such that } r\rho \in \hat{\mathbb{Z}}\}, \tag{3.36}$$

with source and range maps $(r, \rho) \mapsto \rho$ and $(r, \rho) \mapsto r\rho$, and composition

$$(r_1, \rho_1) \circ (r_2, \rho_2) = (r_1 r_2, \rho_2), \quad \text{ if } r_2\rho_2 = \rho_1. \tag{3.37}$$

PROPOSITION 3.22. *The map $\gamma : \mathcal{U}_1 \to \mathcal{G}_1/\mathbb{R}_+^*$ given by*

$$\gamma(r, \rho) = ((r^{-1}\mathbb{Z}, \rho), (\mathbb{Z}, \rho)). \tag{3.38}$$

is an isomorphism of locally compact étale groupoids between $\mathcal{U}_1$ and $\mathcal{G}_1/\mathbb{R}_+^*$.

PROOF. The condition that $r\rho \in \hat{\mathbb{Z}}$ in (3.36) ensures that $(\mathbb{Z}, r\rho)$ is a $\mathbb{Q}$-lattice. Thus so is $(r^{-1}\mathbb{Z}, \rho)$. Moreover, the latter is commensurable to $(\mathbb{Z}, \rho)$ by construction. This shows that the map γ is well defined. Since we know by Lemma 3.21 that 1-dimensional $\mathbb{Q}$-lattices can always be written in the form

$$(\Lambda, \phi) = \lambda(\mathbb{Z}, \rho), \tag{3.39}$$

for some $\lambda \in \mathbb{R}_+^*$ and $\rho \in \hat{\mathbb{Z}}$, we see that the map γ gives an identification between the units $\mathcal{U}_1^{(0)} = \hat{\mathbb{Z}}$ and the set of units $(\mathcal{G}_1/\mathbb{R}_+^*)^{(0)}$, that is, the quotient of the space of 1-dimensional $\mathbb{Q}$-lattices by the scaling action of $\mathbb{R}_+^*$. Moreover, one has

$$\gamma(r_1, \rho_1) \circ \gamma(r_2, \rho_2) = \gamma(r_1 r_2, \rho_2), \quad \text{ if } r_2\rho_2 = \rho_1. \tag{3.40}$$

Indeed using equivalence modulo scaling one gets

$$\gamma(r_2, \rho_2) \sim ((\mathbb{Z}, r_2\rho_2), (r_2\mathbb{Z}, r_2\rho_2)) = ((\mathbb{Z}, \rho_1), (r_2\mathbb{Z}, \rho_1))$$

and

$$\begin{aligned}\gamma(r_1, \rho_1) \circ \gamma(r_2, \rho_2) &\sim ((r_1^{-1}\mathbb{Z}, \rho_1), (\mathbb{Z}, \rho_1)) \circ ((\mathbb{Z}, \rho_1), (r_2\mathbb{Z}, \rho_1)) \\ &= ((r_1^{-1}\mathbb{Z}, \rho_1), (r_2\mathbb{Z}, \rho_1))\end{aligned}$$

which is equivalent modulo scaling to

$$((r_2^{-1}r_1^{-1}\mathbb{Z}, \rho_2), (\mathbb{Z}, \rho_2)) = \gamma(r_1 r_2, \rho_2).$$

Finally, let us show that the map γ is surjective. In fact, up to scaling, every element in $\mathcal{G}_1/\mathbb{R}_+^*$ is of the form

$$((r^{-1}\mathbb{Z}, r^{-1}\rho'), (\mathbb{Z}, \rho)), \tag{3.41}$$

with $\rho, \rho' \in \hat{\mathbb{Z}}$ and $r = \frac{a}{b} \in \mathbb{Q}_+^*$. Since $r^{-1}\rho' = \rho$ modulo $\frac{1}{a}\mathbb{Z}$, one gets $a\rho - b\rho' = 0$ and $r^{-1}\rho' = \rho$. This completes the proof that γ is a groupoid isomorphism. ☐

To a locally compact étale groupoid $\mathcal{G}$ one can associate a corresponding groupoid C^*-algebra $C^*(\mathcal{G})$, obtained as a completion in the norm (3.44) of the algebra of compactly supported functions on $\mathcal{G}$ with the convolution product

$$f_1 * f_2(g) = \sum_{g_1 g_2 = g} f_1(g_1) f_2(g_2), \tag{3.42}$$

and involution $f^*(g) = \overline{f(g^{-1})}$. For an étale groupoid $\mathcal{G}$, every unit $y \in \mathcal{G}^{(0)}$ defines a representation π_y by left convolution of the algebra $\mathcal{A}_c(\mathcal{G})$ on the Hilbert space

$\mathcal{H}_y = \ell^2(\mathcal{G}_y)$, where $\mathcal{G}_y$ denotes the set of elements in $\mathcal{G}$ with source y. Namely, one has

$$(\pi_y(f)\xi)(g) = \sum_{g_1 g_2 = g} f(g_1)\xi(g_2). \tag{3.43}$$

Suppose $\mathcal{G}^{(0)}$ is compact. Then the algebra is unital. One can then complete $\mathcal{A}_c(\mathcal{G})$ to a C^*-algebra $C^*(\mathcal{G})$ by considering the norm

$$\|f\| := \sup_{y \in \mathcal{G}^{(0)}} \|\pi_y(f)\|_{\mathcal{B}(\mathcal{H}_y)}. \tag{3.44}$$

In the case of the groupoid $\mathcal{G}$ of an equivalence relation $\sim$ on a space X, the algebra is obtained by considering the convolution algebra $\mathcal{A}_c(\mathcal{G})$ of compactly supported functions on the graph of the equivalence relation in $X \times X$, with the convolution product

$$f_1 * f_2(x, x') = \sum_{x \sim u \sim x'} f_1(x, u) f_2(u, x') \tag{3.45}$$

and the involution $f^*(x, x') = \overline{f(x', x)}$.
In the case of the groupoid$\mathcal{G}_1/\mathbb{R}^*_+$ of the equivalence relation of commensurability on 1-dimensional $\mathbb{Q}$-lattices up to scaling, using the equivalent description of Proposition 3.22 in terms of the groupoid$\mathcal{U}_1$, we obtain the following description of the corresponding groupoid C^*-algebra (cf. [**30**], [**191**], [**192**]).

Proposition 3.23. *The groupoid C^*-algebra $\mathcal{A}_1 := C^*(\mathcal{G}_1/\mathbb{R}^*_+)$ is isomorphic to the semigroup crossed product $C(\hat{\mathbb{Z}}) \rtimes \mathbb{N}$, with $\mathbb{N} = \mathbb{Z}_{>0}$. This has an explicit presentation with generators μ_n, $n \in \mathbb{N}$ and $e(r)$, for $r \in \mathbb{Q}/\mathbb{Z}$, satisfying the relations*

- $\mu_n^* \mu_n = 1$, *for all $n \in \mathbb{N}$,*
- $\mu_k \mu_n = \mu_{kn}$, *for all $k, n \in \mathbb{N}$,*
- $e(0) = 1$, $e(r)^* = e(-r)$, *and* $e(r)e(s) = e(r+s)$, *for all $r, s \in \mathbb{Q}/\mathbb{Z}$,*
- *For all $n \in \mathbb{N}$ and all $r \in \mathbb{Q}/\mathbb{Z}$,*

$$\mu_n \, e(r) \, \mu_n^* = \frac{1}{n} \sum_{ns = r} e(s). \tag{3.46}$$

The algebra $\mathcal{A}_1$ is Morita equivalent to the crossed product algebra

$$C_0(\mathbb{A}_{\mathbb{Q},f}) \rtimes \mathbb{Q}^*_+. \tag{3.47}$$

Proof. For each integer $n \in \mathbb{N} = \mathbb{Z}_{>0}$, we let $n\hat{\mathbb{Z}} \subset \hat{\mathbb{Z}}$ be the range of multiplication by n. It is an open and closed subset of $\hat{\mathbb{Z}}$, whose characteristic function π_n is a projection $\pi_n \in C(\hat{\mathbb{Z}})$. By construction, the π_n satisfy

$$\pi_n \, \pi_m = \pi_{n \vee m}, \quad \forall n, \, m \in \mathbb{N}, \tag{3.48}$$

where $n \vee m$ denotes the lowest common multiple of n and m. In terms of functions of $\mathbb{Q}$-lattices, π_n can be identified with the characteristic function of the set of 1-dimensional $\mathbb{Q}$-lattices that are divisible by n in the sense of Definition 3.20.
The semigroup action of $\mathbb{N}$ on $C(\hat{\mathbb{Z}})$ corresponds to the isomorphism

$$\alpha_n(f)(\rho) := f(n^{-1}\rho), \quad \forall \rho \in n\hat{\mathbb{Z}} \tag{3.49}$$

of $C(\hat{\mathbb{Z}})$ with the reduced algebra $C(\hat{\mathbb{Z}})_{\pi_n}$ of $C(\hat{\mathbb{Z}})$ by the projection π_n.

Equivalently, in terms of $\mathbb{Q}$-lattices, the semigroup action is given by

$$\alpha_n(f)(\Lambda,\phi) = f(n\Lambda,\phi), \tag{3.50}$$

when (Λ,ϕ) is in the support of π_n and $\alpha_n(f)(\Lambda,\phi) = 0$ otherwise.
Let μ_n be the function $\mu_n(r,\rho)$ that vanishes unless $r = n$ and is equal to 1 for $r = n$. These satisfy the relations $\mu_n^*\mu_n = 1$ for all n, $\mu_k\mu_n = \mu_{kn}$, for all k,n, and $\pi_n = \mu_n\mu_n^*$. The semigroup action is implemented by

$$\mu_n f \mu_n^* = \alpha_n(f), \quad \forall f \in C(\hat{\mathbb{Z}}). \tag{3.51}$$

Notice that the inclusion $\mathbb{Q} \subset \mathbb{A}_{\mathbb{Q},f}$ induces an identification $\mathbb{Q}/\mathbb{Z} = \mathbb{A}_{\mathbb{Q},f}/\hat{\mathbb{Z}}$. Using this identification one gets a natural character $\mathbf{e}$ of $\mathbb{A}_{\mathbb{Q},f}/\hat{\mathbb{Z}}$ such that

$$\mathbf{e}(r) = \exp(2\pi i r), \quad \forall r \in \mathbb{Q}/\mathbb{Z}, \tag{3.52}$$

and a pairing of $\mathbb{Q}/\mathbb{Z}$ with $\hat{\mathbb{Z}}$ of the form

$$\langle r, x\rangle = \mathbf{e}(rx), \quad \forall r \in \mathbb{Q}/\mathbb{Z}, \forall x \in \hat{\mathbb{Z}}. \tag{3.53}$$

The locally compact abelian group $\mathbb{A}_{\mathbb{Q},f}$ is its own Pontrjagin dual using the pairing $\langle x, y\rangle = \mathbf{e}(xy)$ (cf. [**293**]). Thus using the pairing (3.53) we can identify the group $\mathbb{Q}/\mathbb{Z}$ with the Pontrjagin dual of $\hat{\mathbb{Z}}$. By Pontrjagin duality, we have the identification $C(\hat{\mathbb{Z}}) = C^*(\mathbb{Q}/\mathbb{Z})$. The explicit presentation is then obtained using this identification and the crossed product relation (3.51) implemented by the isometries μ_n, cf. [**30**], [**191**].
The Morita equivalence follows from the fact that the algebra $C(\hat{\mathbb{Z}}) \rtimes \mathbb{N}$ is a full corner of $C_0(\mathbb{A}_{\mathbb{Q},f}) \rtimes \mathbb{Q}_+^*$, namely

$$C(\hat{\mathbb{Z}}) \rtimes \mathbb{N} = (C_0(\mathbb{A}_{\mathbb{Q},f}) \rtimes \mathbb{Q}_+^*)_\pi, \tag{3.54}$$

where the right hand side is the algebra obtained by compression of $C_0(\mathbb{A}_{\mathbb{Q},f}) \rtimes \mathbb{Q}_+^*$ with the projection π given by the characteristic function of $\hat{\mathbb{Z}} \subset \mathbb{A}_{\mathbb{Q},f}$, as shown in [**192**]. ◻

The algebra $\mathcal{A}_1$ is the noncommutative algebra of coordinates of the noncommutative space of commensurability classes of $\mathbb{Q}$-lattices up to scale, according to the general idea that quotients of equivalence relations are best described in terms of noncommutative spaces using the (noncommutative) convolution algebra of functions on the graph of the equivalence relation as the "ring of functions on the quotient". We are going to give a more geometric description of the underlying noncommutative space in §4.8 below in terms of noncommutative Shimura varieties.

4.1. The Bost–Connes system.

We now show that the noncommutative space of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices up to scale carries a natural time evolution which makes $\mathcal{A}_1$ the algebra of observables of a quantum statistical mechanical system. We see in the next section that the system obtained in this way is the same as the one introduced by Bost and Connes in [**30**].

As above, let $\mathcal{A}_1$ be the C^*-algebra $C^*(\mathcal{G}_1/\mathbb{R}_+^*)$ of the quotient by the scaling action of the groupoidof the commensurability relation on the space of 1-dimensional $\mathbb{Q}$-lattices. We can describe elements in $\mathcal{A}_1$ as functions of pairs of commensurable

$\mathbb{Q}$-lattices

$$f((\Lambda,\phi),(\Lambda',\phi')), \quad (\Lambda,\phi)\sim(\Lambda',\phi'), \tag{3.55}$$

which are homogeneous of degree zero with respect to the scaling action, namely

$$f(\lambda(\Lambda,\phi),\lambda(\Lambda',\phi')) = f((\Lambda,\phi),(\Lambda',\phi')), \quad \forall \lambda\in\mathbb{R}_+^*. \tag{3.56}$$

The time evolution $\sigma:\mathbb{R}\to \mathrm{Aut}(\mathcal{A}_1)$ is determined by the covolume of a pair of commensurable $\mathbb{Q}$-lattices. Namely, for $L=(\Lambda,\phi)$ and $L'=(\Lambda',\phi')$ a pair of commensurable $\mathbb{Q}$-lattices $L\sim L'$, let $|L/L'|$ denote the ratio of the covolumes

$$|L/L'| = \frac{\mathrm{Covol}(\Lambda')}{\mathrm{Covol}(\Lambda)}. \tag{3.57}$$

This is clearly an invariant under scaling of both lattices,

$$\frac{\mathrm{Covol}(\lambda\Lambda')}{\mathrm{Covol}(\lambda\Lambda)} = \frac{\mathrm{Covol}(\Lambda')}{\mathrm{Covol}(\Lambda)}.$$

Moreover when $L'\subset L$, the order of the quotient group L/L' is equal to $|L/L'|$. The time evolution on the algebra $\mathcal{A}_1$ is given by

$$\sigma_t(f)(L,L') := |L/L'|^{it} f(L,L'). \tag{3.58}$$

LEMMA 3.24. *In terms of the equivalent description of $\mathcal{A}_1$ as $C^*(\mathcal{U}_1)$, the time evolution* (3.58) *is of the form*

$$\sigma_t(f)(r,\rho) = r^{it} f(r,\rho), \quad \forall (r,\rho)\in\mathcal{U}_1. \tag{3.59}$$

In terms of the explicit presentation of Proposition 3.23, this time evolution is of the form

$$\sigma_t(\mu_n) = n^{it}\mu_n, \qquad \sigma_t(e(r)) = e(r). \tag{3.60}$$

PROOF. The result follows easily from Propositions 3.22 and 3.23. By (3.38) one has

$$\gamma(r,\rho) = ((r^{-1}\mathbb{Z},\rho),(\mathbb{Z},\rho)),$$

while

$$|(r^{-1}\mathbb{Z})/\mathbb{Z}| = \frac{\mathrm{Covol}(\mathbb{Z})}{\mathrm{Covol}((r^{-1}\mathbb{Z}))} = r$$

□

We now show that the quantum statistical mechanical system $(\mathcal{A}_1,\sigma_t)$ is the same as the Bost–Connes system of [**29**], [**30**].

4.2. Hecke algebras.

The original construction of the Bost–Connes (BC) quantum statistical mechanical system is based on Hecke algebras of quasi-normal pairs. The main conceptual steps involved in the construction of the original BC system are the following.

- The construction, due to Hecke, of convolution algebras associated to double cosets of algebraic groups over the rational numbers.
- The existence of a canonical time evolution on a noncommutative von Neumann algebra.

One considers here the "$ax + b$" algebraic group P, i.e. the functor which to any abelian ring R assigns the group P_R of 2 by 2 matrices over R of the form

$$P_R = \left\{ \begin{pmatrix} 1 & b \\ 0 & a \end{pmatrix} \; ; \; a, b \in R \, , \; a \text{ invertible} \right\} . \tag{3.61}$$

By construction $P_{\mathbb{Z}}^{+} \subset P_{\mathbb{Q}}^{+}$ is an inclusion $\Gamma_0 \subset \Gamma$ of countable groups, where $P_{\mathbb{Z}}^{+}$ and $P_{\mathbb{Q}}^{+}$ denote the restrictions to $a > 0$. This inclusion fulfills the commensurability condition

(3.62) The orbits of the left action of Γ_0 on Γ/Γ_0 are all *finite*.

The same clearly holds for orbits of Γ_0 acting on the right on $\Gamma_0 \backslash \Gamma$.
The Hecke algebra $\mathcal{H}_{\mathbb{Q}}(\Gamma, \Gamma_0)$ is by definition the convolution algebra of functions of finite support

$$f : \Gamma_0 \backslash \Gamma \to \mathbb{Q}, \tag{3.63}$$

which fulfill the Γ_0-invariance condition

$$f(\gamma\gamma_0) = f(\gamma), \qquad \forall \gamma \in \Gamma, \forall \gamma_0 \in \Gamma_0, \tag{3.64}$$

so that f is really defined on $\Gamma_0 \backslash \Gamma / \Gamma_0$. The convolution product is then given by

$$(f_1 * f_2)(\gamma) = \sum_{\Gamma_0 \backslash \Gamma} f_1(\gamma\gamma_1^{-1}) f_2(\gamma_1) \, . \tag{3.65}$$

The time evolution appears from the analysis of the *regular representation* of the pair (Γ, Γ_0). It is trivial when Γ_0 is normal, or for the original case considered by Hecke of the inclusion $\mathrm{GL}_2(\mathbb{Z}) \subset \mathrm{GL}_2(\mathbb{Q})$, but it becomes interesting in the parabolic case, due to the lack of unimodularity of the parabolic group.
The regular representation

$$(\pi(f)\xi)(\gamma) = \sum_{\Gamma_0 \backslash \Gamma} f(\gamma\gamma_1^{-1}) \xi(\gamma_1) \tag{3.66}$$

of the Hecke algebra $\mathcal{H}_{\mathbb{Q}}(\Gamma, \Gamma_0)$ on the Hilbert space

$$\mathcal{H} = \ell^2(\Gamma_0 \backslash \Gamma) \tag{3.67}$$

extends to the complexification

$$\mathcal{H}_{\mathbb{C}}(\Gamma, \Gamma_0) = \mathcal{H}_{\mathbb{Q}}(\Gamma, \Gamma_0) \otimes_{\mathbb{Q}} \mathbb{C} \tag{3.68}$$

of the Hecke algebra. This inherits from the representation an involution $a \mapsto a^*$, uniquely defined so that $\pi(a^*) = \pi(a)^*$ (the Hilbert space adjoint), namely

$$f^*(\gamma) := \overline{f(\gamma^{-1})}, \quad \forall \gamma \in \Gamma_0 \backslash \Gamma / \Gamma_0. \tag{3.69}$$

By Tomita's theory (cf. [**279**]), the von Neumann algebra generated by $\mathcal{H}_{\mathbb{C}}(\Gamma, \Gamma_0)$ in the regular representation carries a natural time evolution σ_t associated to the state φ_1 given by

$$\varphi_1(f) = \langle \varepsilon_e, \pi(f) \varepsilon_e \rangle \tag{3.70}$$

where ε_e is the cyclic and separating vector for the regular representation given by the left coset $\{\Gamma_0\} \in \Gamma_0 \backslash \Gamma$. It turns out that this time evolution preserves the dense subalgebra $\mathcal{H}_{\mathbb{C}}(\Gamma,\Gamma_0)$ and has the explicit form (cf. [**30**])

$$\sigma_t(f)(\gamma) = \left(\frac{L(\gamma)}{R(\gamma)}\right)^{-it} f(\gamma), \quad \forall \gamma \in \Gamma_0\backslash\Gamma/\Gamma_0\,, \tag{3.71}$$

where the integer valued functions L and R on the double coset space are given respectively by

$$L(\gamma) = \text{ cardinality of left } \Gamma_0 \text{ orbit of } \gamma \text{ in } \Gamma/\Gamma_0\,, \quad R(\gamma) = L(\gamma^{-1})\,. \tag{3.72}$$

Again by Tomita's theory, for this uniquely determined time evolution $\sigma_t \in \mathrm{Aut}(\mathcal{H}_{\mathbb{C}}(\Gamma,\Gamma_0))$, the state φ_1 is a KMS$_\beta$ state at inverse temperature $\beta = 1$.

PROPOSITION 3.25. *The C^*-algebra completion of the Hecke algebra $\mathcal{H}_{\mathbb{C}}(\Gamma,\Gamma_0)$ of the BC system is isomorphic to the C^*-algebra $\mathcal{A}_1$ of Proposition 3.23. The Hecke algebra $\mathcal{H}_{\mathbb{Q}}(\Gamma,\Gamma_0)$ is isomorphic to the $\mathbb{Q}$-algebra generated by the $e(r)$ and μ_n with the relations of Proposition 3.23. The time evolution σ_t of* (3.71) *agrees with the time evolution* (3.58).

PROOF. It is shown in [**30**] that the C^*-completion of the Hecke algebra $\mathcal{H}_{\mathbb{C}}(\Gamma,\Gamma_0) = \mathcal{H}_{\mathbb{Q}}(\Gamma,\Gamma_0)\otimes_{\mathbb{Q}}\mathbb{C}$ has a presentation with generators and relations as in Proposition 3.23 (cf. also [**191**] for the form of the relations). It is also shown in [**30**] that in terms of this presentation the time evolution (3.71) has the form (3.60). ☐

4.3. Symmetries of the BC system.

The group $\hat{\mathbb{Z}}^*$ of invertible elements in $\hat{\mathbb{Z}}$ acts by automorphisms of the algebra $\mathcal{A}_1$, compatibly with the time evolution. In fact, the action is given by

$$g(f)(r,\rho) = f(r,g\rho), \quad \forall g \in \hat{\mathbb{Z}}^*, \forall (r,\rho) \in \mathcal{U}_1. \tag{3.73}$$

In fact, if $(r,\rho)\in\mathcal{U}_1$, then also $(r,g\rho)\in\mathcal{U}_1$. The action (3.73) clearly satisfies the condition $g\sigma_t = \sigma_t g$.

At the level of $\mathbb{Q}$-lattices, the action (3.73) is the same (the group $\hat{\mathbb{Z}}^*$ being commutative) as the action

$$g(\Lambda,\phi) = (\Lambda, \phi\circ g), \tag{3.74}$$

given by precomposing the map $\phi : \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}\Lambda/\Lambda$ with $g \in \mathrm{GL}_1(\hat{\mathbb{Z}})$ viewed as an invertible homomorphism $g : \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$.

In terms of the presentation of the algebra $\mathcal{A}_1$ with generators $e(r)$ and μ_n, this action is of the form

$$g\,\mu_n = \mu_n, \quad g\,e(r) = e(g(r)), \quad \forall g \in \hat{\mathbb{Z}}^*, \forall n \in \mathbb{N}, r \in \mathbb{Q}/\mathbb{Z}, \tag{3.75}$$

where we have used the identification $\hat{\mathbb{Z}} = \mathrm{Hom}(\mathbb{Q}/\mathbb{Z},\mathbb{Q}/\mathbb{Z})$.

Using the formulation of the BC system in terms of the Hecke algebras, one can also see the appearance of the adelic group of symmetries

$$\mathbb{A}^*_{\mathbb{Q},f}/\mathbb{Q}^*_+ = \hat{\mathbb{Z}}^* \tag{3.76}$$

in the following way. The inclusion $\mathbb{Q}\subset\mathbb{A}_{\mathbb{Q},f}$ induces an isomorphism of abelian groups

$$\mathbb{Q}/\mathbb{Z} = \mathbb{A}_{\mathbb{Q},f}/\hat{\mathbb{Z}}, \tag{3.77}$$

where $\hat{\mathbb{Z}}$ is the maximal compact subring of $\mathbb{A}_{\mathbb{Q},f}$. The Hecke algebra $\mathcal{H}_{\mathbb{Q}}(\Gamma,\Gamma_0)$ admits an involutive automorphism α, with $\alpha^2=1$, whose fixed point algebra is the

Hecke algebra of the pair $P_{\mathbb{Z}} \subset P_{\mathbb{Q}}$. This admits an equivalent formulation as Hecke algebra of the pair $P_{\hat{\mathbb{Z}}} \subset P_{\mathbb{A}_{\mathbb{Q},f}}$. This adelic description of the Hecke algebra then displays as a natural symmetry group the idèle class group (3.76), compatibly with the time evolution.

4.4. The arithmetic subalgebra.

We have seen in §4.2 above that there is a natural arithmetic algebra associated to the BC system, namely the Hecke algebra $\mathcal{H}_{\mathbb{Q}}(P^+_{\mathbb{Q}}, P^+_{\mathbb{Z}})$. This is the algebra over $\mathbb{Q}$ generated by the μ_n and the $e(r)$, with the relations of Proposition 3.23. We will denote this algebra in the following as $\mathcal{A}_{1,\mathbb{Q}}$.

In this section we give a description of the algebra $\mathcal{A}_{1,\mathbb{Q}}$ which is more natural from the point of view of $\mathbb{Q}$-lattices and which will lead to the correct generalization in the case of 2-dimensional $\mathbb{Q}$-lattices discussed later in this chapter. The formulation is based on the trigonometric analog of the Eisenstein series, following the analogy between trigonometric and elliptic functions as developed by A. Weil in [**297**].

DEFINITION 3.26. *A function f on the space of 1-dimensional $\mathbb{Q}$-lattices is homogeneous of weight k if it satisfies*

$$f(\lambda(\Lambda,\phi)) = \lambda^{-k} f(\Lambda,\phi), \qquad \forall \lambda \in \mathbb{R}^*_+. \tag{3.78}$$

The description of $\mathbb{Q}$-lattices up to scaling given in Lemma 3.21 implies that we can regard the algebra $C(\hat{\mathbb{Z}})$ as the algebra of functions on the space of 1-dimensional $\mathbb{Q}$-lattices which are homogeneous of weight zero with respect to the scaling action of $\mathbb{R}^*_+$.

We define the function $c(\Lambda)$ to be the multiple of the covolume $|\Lambda|$ of the lattice, specified by the choice of normalization

$$2\pi i\, c(\mathbb{Z}) = 1. \tag{3.79}$$

The function c is homogeneous of weight -1 on the space of 1-dimensional $\mathbb{Q}$-lattices.

We then define, for $a \in \mathbb{Q}/\mathbb{Z}$, a function $e_{1,a}$ of weight zero by setting

$$e_{1,a}(\Lambda,\phi) = c(\Lambda) \sum_{y\in\Lambda+\phi(a)} y^{-1}. \tag{3.80}$$

Here we use the Eisenstein summation convention, namely we take the summation to mean

$$\lim_{N\to\infty} \sum_{-N}^{N}$$

when $\phi(a) \neq 0$. We also set $e_{1,a}(\Lambda,\phi) = 0$ when $\phi(a) = 0$.

For $x \neq 0$ one has, using the above convention,

$$\sum_{y\in\mathbb{Z}+x} y^{-1} = \frac{1}{x} + \sum_{1}^{\infty}\Big(\frac{1}{x+n} + \frac{1}{x-n}\Big) = \pi \cot \pi x \tag{3.81}$$

Thus for $\rho \in \hat{\mathbb{Z}} = \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, \mathbb{Q}/\mathbb{Z})$ and assuming $\rho(a) \neq 0$ one has

$$e_{1,a}(\mathbb{Z},\rho) = \frac{1}{2i} \cot(\pi\rho(a))\,. \tag{3.82}$$

More generally, we define, for each weight $k \in \mathbb{N}$ and each $a \in \mathbb{Q}/\mathbb{Z}$, a function $\epsilon_{k,a}$ on the space of 1-dimensional $\mathbb{Q}$-lattices by setting

$$\epsilon_{k,a}(\Lambda, \phi) = \sum_{y \in \Lambda + \phi(a)} y^{-k} . \tag{3.83}$$

This is well defined, provided $\phi(a) \neq 0$. In the case $\phi(a) = 0$ we set

$$\epsilon_{k,a}(\Lambda, \phi) = \lambda_k c(\Lambda)^{-k}, \tag{3.84}$$

where the constants λ_k are given by

$$\lambda_k = (2^k - 1)\gamma_k, \tag{3.85}$$

with

$$\gamma_k = (2\pi i)^{-k} \sum_{y \in \mathbb{Z}^*} y^{-k} \tag{3.86}$$

One has $\gamma_{2j+1} = 0$ for all j and

$$\gamma_{2j} = (-1)^j \frac{B_j}{(2j)!}, \tag{3.87}$$

with $B_j \in \mathbb{Q}$ the Bernoulli numbers. Equivalently, the γ_{2j} are defined through the generating series

$$\frac{x}{e^x - 1} = 1 - \frac{x}{2} - \sum_{j=1}^{\infty} \gamma_{2j} x^{2j}. \tag{3.88}$$

The function $\epsilon_{k,a}$ has weight k, i.e. it satisfies the homogeneity condition

$$\epsilon_{k,a}(\lambda(\Lambda, \phi)) = \lambda^{-k} \epsilon_{k,a}(\Lambda, \phi), \quad \forall \lambda \in \mathbb{R}_+^*. \tag{3.89}$$

For $a = \frac{b}{N}$, the function $\epsilon_{k,a}$ has level N, in the sense that it only uses the restriction

$$\phi_N : \frac{1}{N}\mathbb{Z}/\mathbb{Z} \to \frac{1}{N}\Lambda/\Lambda$$

of ϕ to N-torsion points of $\mathbb{Q}/\mathbb{Z}$.

We consider the associated weight zero functions $e_{k,a}$ obtained by setting

$$e_{k,a} := c^k \, \epsilon_{k,a}, \tag{3.90}$$

where c is the normalized covolume function of (3.79).

The functions $e_{k,a}$ satisfy the following multiplicative relations, that express the $e_{k,a}$ as polynomial functions of the $e_{1,a}$ of (3.80).

LEMMA 3.27. *The functions $e_{k,a}$ satisfy the relations*

$$e_{k,a} = P_k(e_{1,a}), \tag{3.91}$$

where the P_k are polynomials with rational coefficients uniquely determined by the identities

$$P_1(u) = u, \qquad P_{k+1}(u) = \frac{1}{k}\left(u^2 - \frac{1}{4}\right) \partial_u P_k(u). \tag{3.92}$$

PROOF. For $\phi(a) \notin \Lambda$, the result follows from the elementary formulas for the trigonometric analog of the Eisenstein series (cf. [**297**] Chapter II). Indeed, with $x = \rho(a) \neq 0$ one has

$$e_{k,a}(\mathbb{Z}, \rho) = (2\pi i)^{-k} \sum_{y \in \mathbb{Z}+x} y^{-k} = \sigma_k(x) \tag{3.93}$$

and by construction one has

$$\sigma_{k+1}(x) = -\frac{1}{2\pi i k} \frac{d}{dx} \sigma_k(x) \tag{3.94}$$

while with $u(x) = \sigma_1(x) = \frac{1}{2i} \cot \pi x$ one has

$$-\frac{1}{2\pi i} \frac{du}{dx} = u^2 - \frac{1}{4}\,. \tag{3.95}$$

Thus, using (3.94) one obtains by induction on k the identity

$$\sigma_k = P_k(\sigma_1)\,, \tag{3.96}$$

which proves (3.91) for $\phi(a) \notin \Lambda$.
Since we have set $e_{1,a}(\Lambda, \phi) = 0$ for $\phi(a) \in \Lambda$, the validity of (3.91) in that case uniquely dictates the choice of the normalization constants λ_k to be, using $\sigma_1(\frac{1}{2}) = 0$,

$$\lambda_k = P_k(0) = \sigma_k(\frac{1}{2})\,, \tag{3.97}$$

which vanishes for k odd and is a *non-zero* rational number for k even. In fact one has

$$\sigma_k(\frac{1}{2}) = (2\pi i)^{-k} \sum_{y \in \mathbb{Z}+\frac{1}{2}} y^{-k} = (2\pi i)^{-k}(2^k - 1) \sum_{y \in \mathbb{Z}^*} y^{-k}$$

which gives

$$\lambda_k = (2^k - 1)\gamma_k, \tag{3.98}$$

with γ_k as in (3.86). $\square$

We can express the functions $e_{k,a}$ in terms of the generators $e(r)$ of the algebra $\mathcal{A}_1$. In fact, we can view the $e(r)$ as the functions on $\mathbb{Q}$-lattices

$$e(r)(\Lambda, \phi) := \exp(2\pi i \rho(r))\,, \qquad \text{for } (\Lambda, \phi) = (\lambda\mathbb{Z}, \lambda\rho), \tag{3.99}$$

for $r \in \mathbb{Q}/\mathbb{Z}$ and $\rho \in \hat{\mathbb{Z}} = \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, \mathbb{Q}/\mathbb{Z})$. These are clearly of weight zero in $\lambda \in \mathbb{R}_+^*$.
We have the following relations.

LEMMA 3.28. *Let $a \in \mathbb{Q}/\mathbb{Z}$, and $n > 0$ such that $na = 0$. Then the following holds:*

$$e_{1,a} = \sum_{k=1}^{n-1} \left(\frac{k}{n} - \frac{1}{2}\right) e(ka). \tag{3.100}$$

PROOF. We evaluate both sides on a $\mathbb{Q}$-lattice $(\Lambda, \phi) = (\lambda\mathbb{Z}, \lambda\rho)$, with $\lambda \in \mathbb{R}_+^*$ and $\rho \in \hat{\mathbb{Z}} = \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, \mathbb{Q}/\mathbb{Z})$. Both sides only depend on the restriction $x \mapsto cx$ of

ρ to the n-torsion elements of $\mathbb{Q}/\mathbb{Z}$, which we write as multiplication by $c \in \mathbb{Z}/n\mathbb{Z}$. Let $a = \frac{b}{n}$. If $bc = 0 \mod n$, then $\phi(a) = 0$ and both sides vanish since

$$e(ka)(\Lambda, \phi) = \exp\left(2\pi i \frac{kbc}{n}\right) = 1$$

for all k. If $bc = 0 \mod n$, then $\phi(a) \neq 0$ and, using (3.82) the left hand side is given by

$$\frac{1}{2}\frac{U+1}{U-1}, \tag{3.101}$$

where

$$U = \exp\left(2\pi i \frac{bc}{n}\right),$$

with $U^n = 1$ and $U \neq 1$. The right hand side is given by

$$\sum_{k=1}^{n-1}\left(\frac{k}{n} - \frac{1}{2}\right)U^k,$$

which gives $\frac{1}{2}(U+1)$ after multiplication by $U-1$. Indeed using the identity (Abel summation)

$$\sum_{k=1}^{n-1} a_k(b_{k+1} - b_k) = a_{n-1}b_n - a_1 b_1 + \sum_{k=1}^{n-2}(a_k - a_{k+1})b_{k+1}$$

one has

$$\sum_{k=1}^{n-1} k(U^{k+1} - U^k) = (n-1) - U - \sum_{k=1}^{n-2} U^{k+1} = n$$

while $\sum_{k=1}^{n-1}(U^{k+1} - U^k) = 1 - U$ which gives the required result. □

Equality (3.101) shows that $e_{1,a}$ is (one half of) the Cayley transform of $e(a)$, with care taken when $e(a) - 1$ fails to be invertible. In particular, while $e(a)$ is unitary, $e_{1,a}$ is skew-adjoint, namely

$$e_{1,a}^* = -e_{1,a}. \tag{3.102}$$

In addition to the multiplicative relations of Lemma 3.27 above, the functions $e_{k,a}$ also satisfy division relations. In order to derive these relations, we first need to discuss some divisibility properties of $\mathbb{Q}$-lattices, where divisibility for $\mathbb{Q}$-lattices is defined as in Definition 3.20 above.

As before, we let π_n denote the characteristic function of the set of 1-dimensional $\mathbb{Q}$-lattices (up to scale) that are divisible by n. It corresponds to the characteristic function of $n\hat{\mathbb{Z}} \subset \hat{\mathbb{Z}}$.

For $N > 0$ let (Λ, ϕ) be a $\mathbb{Q}$-lattice. Let $c \in \mathbb{Z}/N\mathbb{Z}$ be such that $\phi_N(a) = ca$ for all $a \in \mathbb{Z}/N\mathbb{Z}$. The order of the kernel of ϕ_N is $m = \gcd\{N, c\}$. Moreover, a divisor $b|N$ divides (Λ, ϕ) if and only if it divides c (i.e. if and only if multiplication by c is 0 in $\mathbb{Z}/b\mathbb{Z}$). Thus, for any function f on $\mathbb{N} = \mathbb{Z}_{>0}$ one has

$$\sum_{b|N} f(b)\pi_b(\Lambda, \phi) = \sum_{b|\gcd\{N,c\}} f(b). \tag{3.103}$$

This allows one to express any function $f(m)$ of the order $m = \gcd\{N, c\}$ of the kernel of ϕ_N in terms of the projections π_b, where $b|N$. Indeed, the right hand side of (3.103) gives such an explicit formula for

$$f(m) = \sum_{b|m} g(b)$$

for any function g and the Möbius inversion formula

$$g(n) = \sum_{d|n} \mu(d) f(n/d) \tag{3.104}$$

shows that any function f can be written in that way. In order to obtain the function $m \mapsto m^j$ we set

$$f_j(n) := \sum_{d|n} \mu(d)(n/d)^j, \tag{3.105}$$

where μ is the Möbius function , so that

$$f_j(n) = n^j \prod_{p\,\mathrm{prime},\, p|n} (1 - p^{-j}). \tag{3.106}$$

In particular f_1 is the Euler totient function.
The Möbius inversion formula together with (3.103) then give

$$\sum_{b|N} f_j(b) \pi_b(\Lambda, \phi) = m^j, \qquad \text{for } m = \gcd\{N, c\}. \tag{3.107}$$

Thus, we obtain the division relations fulfilled by the functions $e_{k,a}$ as follows.

LEMMA 3.29. *For a given $N > 0$, the functions $e_{k,a}$ satisfy the relations*

$$\sum_{Na=0} e_{k,a} = \gamma_k \sum_{d|N} ((2^k - 2) f_1(d) + N^k f_{-k+1}(d)) \pi_d. \tag{3.108}$$

PROOF. Consider a $\mathbb{Q}$-lattice (Λ, ϕ) such that $\operatorname{Ker} \phi_N$ is of order $m|N$, so that $N = md$. Then the result follows from the identity (using (3.86))

$$\begin{aligned} \sum\nolimits_{Na=0} \epsilon_{k,a}(\Lambda, \phi) &= m \sum\nolimits_{y \in \frac{1}{d}\Lambda/\Lambda} y^{-k} + m(2^k - 1)\gamma_k c^{-k}(\Lambda) \\ &= m(d^k + 2^k - 2)\gamma_k c^{-k}, \end{aligned} \tag{3.109}$$

together with (3.107) applied for $j = 1$ and $j = 1 - k$ using

$$m(d^k + 2^k - 2) = (2^k - 2)m + N^k m^{-k+1}$$

□

We can now derive the main result of this section.

THEOREM 3.30. *The functions $e_{1,a}$ of (3.80) have the following properties.*
1. *The $e_{1,a}$, for $a \in \mathbb{Q}/\mathbb{Z}$, generate the group ring $\mathbb{Q}[\mathbb{Q}/\mathbb{Z}]$.*
2. *The arithmetic algebra $\mathcal{A}_{1,\mathbb{Q}}$ is generated by the $e_{1,a}$, for $a \in \mathbb{Q}/\mathbb{Z}$, and the μ_n and μ_n^*.*
3. *The complexification $\mathcal{A}_{1,\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{C}$ is the subalgebra of the C^*-algebra $\mathcal{A}_1$ generated by these elements.*

PROOF. (1) Let $\mathcal{B}_{\mathbb{Q}}$ denote the algebra generated over $\mathbb{Q}$ by the $e_{1,a}$. By Lemma 3.28 it is a subalgebra of the group ring $\mathbb{Q}[\mathbb{Q}/\mathbb{Z}]$. By lemma 3.27 one has $e_{k,a} \in \mathcal{B}_{\mathbb{Q}}$ for all k.
We first show that the projections π_N belong to $\mathcal{B}_{\mathbb{Q}}$. The idea is to use (3.108) to express π_N as a rational linear combination of the $e_{k,a}$, with k even. Since for relatively prime integers n, m one has $\pi_{nm} = \pi_n \pi_m$, we can assume that $N = p^b$ is a prime power. However, special care is needed when N is a power of two. In fact, when $N = p^b$ is a prime power, the coefficient of $\gamma_k \pi_N$ in (3.108) is given by

$$(2^k - 2) f_1(N) + N^k f_{-k+1}(N) = (2^k - 2)(p-1)p^{b-1} - p^b(p^{k-1} - 1)$$

using (3.106). This does not vanish unless $p = 2$, and is equal to $-p^{b-1}(2 - 3p + p^2)$ for $k = 2$. Thus, one can express π_N as a linear combination of the $e_{2,a}$ by induction on b. For $p = 2$, $N = 2^b$, $b > 1$, the coefficient of $\gamma_k \pi_N$ in (3.108) is zero but the coefficient of $\gamma_k \pi_{N/2}$ is $-2^{b-2}(2^k - 1)(2^k - 2) \neq 0$ for k even. This allows one to still express π_N as a linear combination of the $e_{2,a}$ by induction on b. Thus, for instance, π_2 is given by

$$\pi_2 = 3 + 2 \sum_{4a=0} e_{2,a}.$$

In general, π_{2^n} involves

$$\sum_{2^{n+1}a=0} e_{2,a}.$$

Since for relatively prime integers n, m one has $\pi_{nm} = \pi_n \pi_m$, we see that the algebra $\mathcal{B}_{\mathbb{Q}}$ generated over $\mathbb{Q}$ by the $e_{1,a}$ contains all the projections π_n.
In order to show that $\mathcal{B}_{\mathbb{Q}}$ contains the $e(r)$ it is enough to show that, for any prime power $N = p^b$, it contains $e(\frac{1}{N})$. This is proved by induction on b as follows.
Multiplying the relation (3.108) by $1 - \pi_p$ and using the identity $(1 - \pi_p)\pi_{p^\ell} = 0$ for $\ell > 0$ (which follows from (3.48)), we obtain the equalities

$$(1 - \pi_p) \sum_{Na=0} e_{k,a} = (N^k + 2^k - 2)\gamma_k (1 - \pi_p). \tag{3.110}$$

We then set

$$z(j) := (1 - \pi_p) e_{1, \frac{j}{N}}. \tag{3.111}$$

The relations (3.110) together with (3.91) show that, in the reduced algebra $(\mathcal{B}_{\mathbb{Q}})_{1-\pi_p}$ one has

$$\sum_{j=1}^{N-1} P_k(z(j)) = (N^k - 1)\gamma_k \tag{3.112}$$

for all k (where we used $P_k(0) = (2^k - 1)\gamma_k$ to subtract the contribution of $j = 0$). The symmetric functions of the $z(j)$ (for $j \in \{1, ..., N-1\}$) are fixed rational numbers σ_h. In particular $z = z(1)$ fulfills

$$Q(z) = z^{N-1} + \sum_{h=1}^{N-1} (-1)^h \sigma_h z^{N-1-h} = 0. \tag{3.113}$$

Using (3.82) the roots of this equation are the

$$\frac{1}{2i} \cot\left(\frac{\pi j}{N}\right), \quad j \in \{1, ..., N-1\} \tag{3.114}$$

Thus, $\pm\frac{1}{2}$ is not a root of (3.113) and this allows us, using the companion matrix of Q, to express the Cayley transform of $2z$ as a polynomial with rational coefficients,

$$\frac{2z+1}{2z-1} = \sum_{n=0}^{N-2} u_n z^n. \tag{3.115}$$

One then has

$$\sum_{0}^{N-2} u_n z^n = (1-\pi_p)e(\frac{1}{N}), \tag{3.116}$$

where the left-hand side belongs to $\mathcal{B}_{\mathbb{Q}}$ by construction.
Notice that the semigroup action of $\mathbb{N}$ given by (3.50) preserves the algebra $\mathcal{B}_{\mathbb{Q}}$ generated over $\mathbb{Q}$ by the $e_{1,a}$ for $a \in \mathbb{Q}/\mathbb{Z}$. In fact, one has

$$\alpha_n(e_{k,a}) = \pi_n\, e_{k,a/n}, \tag{3.117}$$

independently of the choice of the solution $x = a/n$ of $nx = a$. The projections π_n belong to $\mathcal{B}_{\mathbb{Q}}$, hence we see that α_n preserves $\mathcal{B}_{\mathbb{Q}}$.
Now we use the fact that $\pi_p\, e(\frac{1}{N})$ is equal to $\alpha_p(e(\frac{p}{N}))$, for α_p as in (3.50). It follows from the induction hypothesis on b (with $N = p^b$) that $e(\frac{p}{N}) \in \mathcal{B}_{\mathbb{Q}}$, hence using (3.117) that $\alpha_p(e(\frac{p}{N})) \in \mathcal{B}_{\mathbb{Q}}$. Thus, we get $\pi_p\, e(\frac{1}{N}) \in \mathcal{B}_{\mathbb{Q}}$ and using (3.116) that $e(\frac{1}{N}) \in \mathcal{B}_{\mathbb{Q}}$ as required. This completes the proof of (1).
(2) The crossed product by $\mathbb{N}$ is obtained by adjoining to the rational group ring of $\mathbb{Q}/\mathbb{Z}$ the isometries μ_n and their adjoints μ_n^* with the relation

$$\mu_n f \mu_n^* = \alpha_n(f), \qquad \forall f \in \mathbb{Q}[\mathbb{Q}/\mathbb{Z}], \tag{3.118}$$

which gives the rational algebra $\mathcal{A}_{1,\mathbb{Q}}$. (3) follows immediately from (2). $\square$

It is important to remark here that the multiplication and division relations of Lemma 3.27 and Lemma 3.29 are not enough to give a full presentation of the algebra $\mathcal{B}_{\mathbb{Q}}$. In fact, there are more elaborate division relations, which we did not need in the proof of Theorem 3.30, but which contribute extra relations between the generators $e_{1,a}$.
In order to formulate these other division relations, we first define projections $\pi(N,d)$, for $d|N$, belonging to the algebra generated by the π_b with $b|N$. The projection $\pi(N,d)$ is given by the characteristic function of the subset of $\mathbb{Q}$-lattices (Λ,ϕ) for which, letting $c \in \mathbb{Z}/N\mathbb{Z}$ be such that $\phi_N(a) = ca$ for all a, one has

$$\gcd\{N,c\} = N/d.$$

Thus, the projection $\pi(N,d)$ is given by

$$\pi(N,d) = \pi_{N/d} \prod_{k|d} (1-\pi_{kN/d}), \tag{3.119}$$

where the product is over nontrivial divisors $k \neq 1$ of d.
We then have the following relations.

PROPOSITION 3.31. *The functions $e_{k,a}$, for $a \in \mathbb{Q}/\mathbb{Z}$ and for k odd, fulfill the relation*

$$\frac{1}{N}\sum_{Na=0} e_{k,x+a} = \sum_{d|N} \pi(N,d) d^{k-1} e_{k,dx}, \tag{3.120}$$

for any $x \in \mathbb{Q}/\mathbb{Z}$ and any positive integer N.

PROOF. Let (Λ, ϕ) be a $\mathbb{Q}$-lattice such that $\phi_N(a) = ca$ for $c \in \mathbb{Z}/N\mathbb{Z}$ with $\gcd\{N, c\} = N/d = m$. We can assume by homogeneity that $\Lambda = \mathbb{Z}$. Then, when a ranges over the torsion points j/N, with $j \in \{0, \ldots, N-1\}$, the values $\phi(a)$ range m times over the j/d, for $j \in \{0, \ldots, d-1\}$. Thus, the left hand side of (3.31) gives m times

$$c(\mathbb{Z})^k \sum_{j=0}^{d-1} \sum_{y \in \mathbb{Z}+\phi(x)+\frac{j}{d}} y^{-k} = c(\mathbb{Z})^k d^k \sum_{y \in \mathbb{Z}+\phi(dx)} y^{-k}. \tag{3.121}$$

This is clear when $y = 0$ does not appear in the sums involved. When it does, one has

$$\sum_{j=0}^{d-1} \sum_{y \in \mathbb{Z}+\phi(x)+\frac{j}{d}} (y+\epsilon)^{-k} = d^k \sum_{y \in \mathbb{Z}+\phi(dx)} (y+d\epsilon)^{-k}, \tag{3.122}$$

for $\epsilon \notin \frac{\mathbb{Z}}{d}$. Subtracting the pole part on both sides and equating the finite values gives the desired equality. In fact, for odd k, the value of $\epsilon_{k,a}(\Lambda, \phi)$ for $\phi(a) = 0$ can be written as the finite part at $\epsilon = 0$ of

$$\sum_{y \in \Lambda+\phi(a)} (y+\epsilon)^{-k}.$$

$\square$

For even k this no longer holds and the finite value $\gamma_k c(\Lambda)^k$ is replaced by $(2^k - 1)\gamma_k c(\Lambda)^k$. Thus, when $\phi(dx) \in \mathbb{Z}$, one gets an additional term which is best taken care of by multiplying the right hand side in Proposition 3.31 by $(1 - \pi_{\delta(dx)})$, with $\delta(y)$ the order of y in $\mathbb{Q}/\mathbb{Z}$, and adding corresponding terms to the formula, which becomes

$$\begin{aligned} \frac{1}{N} \sum_{Na=0} e_{k,x+a} = {} & \sum_{d|N} \pi(N,d)(1 - \pi_{\delta(d\,x)}) d^{k-1} e_{k,dx} \\ & + \gamma_k \sum_{d|N} (d^{k-1} + d^{-1}(2^k - 2)) \pi(N,d) \pi_{\delta(dx)}. \end{aligned} \tag{3.123}$$

These relations are more elaborate than the division relations for trigonometric functions. They restrict to the latter on the subset of invertible 1-dimensional $\mathbb{Q}$-lattices, for which all π_n, $n \neq 1$ are zero and the only non-zero term on the right hand side is the term with $d = N$. The standard discussion of Eisenstein series in higher dimension is restricted to invertible $\mathbb{Q}$-lattices, but in our case the construction of the endomorphisms implemented by the μ_n requires the extension to non-invertible $\mathbb{Q}$-lattices considered above.

4.5. Class field theory and the Kronecker–Weber theorem.

A local field is by definition a field which can be endowed with a non-discrete locally compact topology (compatible with the field structure). A non-Archimedean local field K is the field of fractions of a complete discrete valuation ring $\mathcal{O}$, such that the residue field $k = \mathcal{O}/\mathfrak{m}$ is a finite field, where $\mathfrak{m}$ is the maximal ideal of $\mathcal{O}$. The field K is endowed with the topology defined by the absolute value $|\cdot|_v = \exp(-v(\cdot))$, where $v : K^* \to \mathbb{Z}$ is the valuation. An Archimedean local field is either $\mathbb{R}$ or $\mathbb{C}$. A non-Archimedean local field of characteristic zero is a finite extension of $\mathbb{Q}_p$, while a non-Archimedean local field of positive characteristic $p > 0$ is of the form $\mathbb{F}_q((t))$ for $q = p^r$ for some $r \in \mathbb{N}$.

Local class field theory shows that there exists a homomorphism (the local Artin homomorphism)

$$\theta : K^* \to \mathrm{Gal}(K^{\mathrm{ab}}/K), \tag{3.124}$$

where K^{ab} is the maximal abelian extension of K in K^{sep}, a fixed separable closure. Upon passing to the profinite completion it induces an isomorphism $\hat{K}^* \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$. The profinite completion of a topological group G is in general defined as $\hat{G} = \varprojlim_U G/U$, with U ranging over finite-index open normal subgroups.

When K is Archimedean, then (3.124) is surjective with kernel given by the connected component of the identity in K^*. If K is non-Archimedean then (3.124) is injective. One can write $K^* = \mathcal{O}^* \pi^{\mathbb{Z}}$, where the elements of $\mathcal{O}^*$ are the units of $\mathcal{O}$ and π is a uniformizer, i.e. an element of $\mathcal{O}$ with $v(\pi) = 1$. Then the Artin homomorphism induces an isomorphism $\mathcal{O}^* \times \pi^{\hat{\mathbb{Z}}} \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$. Under this identification one finds $\theta(\mathcal{O}^*) = \mathrm{Gal}(K^{\mathrm{ab}}/K^{\mathrm{un}})$, where K^{un} is the maximal unramified extension of K in K^{sep}. The group $\mathrm{Gal}(K^{\mathrm{ab}}/K^{\mathrm{un}})$ is the inertia subgroup of $\mathrm{Gal}(K^{\mathrm{ab}}/K)$. The group $\pi^{\hat{\mathbb{Z}}}$ is identified with $\mathrm{Gal}(K^{\mathrm{un}}/K) = \mathrm{Gal}(k^{\mathrm{sep}}/k)$, which is generated by the Frobenius. Thus the Artin morphism maps a uniformizer to the Frobenius automorphism in $\mathrm{Gal}(K^{\mathrm{ab}}/K)$.

Global fields of characteristic zero are number fields, i.e. finite algebraic extensions of $\mathbb{Q}$, while global fields of positive characteristic $p > 0$ are function fields, i.e. the field of functions $\mathbb{F}_q(C)$ of a smooth, projective, geometrically connected curve C over a finite field $\mathbb{F}_q$. These are finite extensions of $\mathbb{F}_p(t)$, the field of functions of $\mathbb{P}^1|_{\mathbb{F}_p}$.

The set $\Sigma_{\mathbb{K}}$ of places of a global field $\mathbb{K}$ is the set of uniform structures on $\mathbb{K}$ whose completion gives a local field. Such structures are still called valuations on $\mathbb{K}$ even in the Archimedean case. One denotes by $\mathbb{K}_v$ the local field obtained by taking the completion of $\mathbb{K}$ in the valuation v. For a non-Archimedean valuation v, one denotes by $\mathcal{O}_v$ the valuation subring of $\mathbb{K}_v$.

For $\mathbb{K}$ a global field, one can construct the ring of adèles $\mathbb{A}_{\mathbb{K}}$ (cf. Chapter II §8.2) as the restricted product

$$\mathbb{A}_{\mathbb{K}} = \prod_{v \in \Sigma_{\mathbb{K}}}{}' \mathbb{K}_v, \tag{3.125}$$

namely those elements $(a_v) \in \prod_v \mathbb{K}_v$ such that $a_v \in \mathcal{O}_v$ for all but finitely many $v \in \Sigma_{\mathbb{K}}$. The idèle class group is the quotient $C_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$.

We have already encountered the adèles $\mathbb{A}_{\mathbb{Q}}$ in the previous chapter in connection with the quantum statistical mechanical system that gives the counting of zeros of the Riemann zeta function and in the spectral realization of its zeros obtained in [**71**]. We will also see the specific case of imaginary quadratic fields $\mathbb{K} = \mathbb{Q}(\sqrt{-d})$ later in this chapter.

In the global case one also has a version of the Artin homomorphism. The global Artin homomorphism

$$\theta : C_{\mathbb{K}} \to \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K}) \tag{3.126}$$

induces an isomorphism $\hat{C}_{\mathbb{K}} \to \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K})$.

In particular, if $\mathbb{K}$ is a number field, then (3.126) is surjective and the kernel is the connected component $D_{\mathbb{K}}$ of the identity on $C_{\mathbb{K}}$. Namely, one has the global class field theory isomorphism

$$\theta : C_{\mathbb{K}}/D_{\mathbb{K}} \stackrel{\simeq}{\to} \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K}). \tag{3.127}$$

If $\mathbb{K}$ is a function field with constant field $k = \mathbb{F}_q$, then (3.126) is injective. The image $\theta(C_{\mathbb{K}})$ in $\mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K})$ consists of those elements whose restriction to $\mathrm{Gal}(k^{\mathrm{sep}}/k)$ is an integer power of the Frobenius (as opposed to a $\hat{\mathbb{Z}}$ power).

In fact, when one looks more closely at the Artin homomorphism, one finds that it satisfies a functoriality property under field extensions, in the form of a commutative diagram

$$\begin{tikzcd} C_{\mathbb{L}} \arrow[r, "\theta"] \arrow[d, "\mathfrak{n}_{\mathbb{L}/\mathbb{K}}"'] & \mathrm{Gal}(\mathbb{L}^{\mathrm{ab}}/\mathbb{L}) \arrow[d, "\mathrm{res}"] \\ C_{\mathbb{K}} \arrow[r, "\theta"] & \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K}) \end{tikzcd} \tag{3.128}$$

where $res : \mathrm{Gal}(\mathbb{L}^{\mathrm{ab}}/\mathbb{L}) \to \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K})$ is the restriction map and $\mathfrak{n}_{\mathbb{L}/\mathbb{K}}$ is the norm. Moreover, for a finite abelian extension $\mathbb{L}$ of a number field $\mathbb{K}$, there is a map called the classical Artin homomorphism

$$\theta_{\mathbb{L}/\mathbb{K}} : I^S_{\mathbb{K}} \to \mathrm{Gal}(\mathbb{L}/\mathbb{K}), \tag{3.129}$$

where S is a finite set of primes which includes all the primes of $\mathbb{K}$ that ramify in $\mathbb{L}$ and $I^S_{\mathbb{K}}$ is the group of fractional ideals that do not involve primes in S. For number fields, the Artin map $\theta_{\mathbb{L}/\mathbb{K}}$ is surjective.

These general result of class field theory, however, do not give an explicit way of constructing abelian extensions. The "explicit class field theory problem" (Hilbert's 12th problem) is precisely the problem of giving explicit generators for the maximal abelian extension of a global field, together with the explicit form of the Galois action. It is considered of central importance to number theory.

In the case of number fields, there are only two cases in which a complete solution to the explicit class field theory problem is known. The case where $\mathbb{K} = \mathbb{Q}$ where the Kronecker–Weber theorem shows that the maximal abelian extension $\mathbb{Q}^{\mathrm{ab}}$ is the cyclotomic field $\mathbb{Q}^{\mathrm{cycl}}$ generated by adjoining the roots of unity. More precisely, the Kronecker–Weber theorem proves that any abelian number field is contained in a field of the form $\mathbb{Q}(\zeta)$, with ζ a root of unity. This then shows that it is possible to generate all abelian extensions of $\mathbb{Q}$ by a single transcendental function, the exponential $\exp(2\pi i\cdot)$, evaluated on the points of $\mathbb{Q}/\mathbb{Z}$, the torsion points of a 1-dimensional lattice.

Kronecker's Jugendtraum was the idea that one could indeed generate abelian extensions of number fields by means of the values of certain transcendental functions. This idea had a remarkable success in the other known case for number fields, namely for $\mathbb{K}$ an imaginary quadratic field. In this case, what is central to the construction of abelian extension is the geometry of elliptic curves with complex multiplication and the transcendental functions are the Weierstrass $\wp$ function and its derivative at torsion points of 2-dimensional lattices, which give the coordinates of the torsion points of elliptic curves. We shall return to discuss the case of imaginary quadratic fields in detail later in this chapter.

The case of function fields can also, to some extent be seen as a generalization of this same principle. In fact, one can generate abelian extensions using an analog of the exponential function, which is given by Drinfeld modules of rank one. This case has been studied recently in the context of noncommutative geometry and quantum statistical mechanics in [**172**] and [**104**].

We concentrate now only on the Kronecker–Weber case, which is the one relevant to the Bost–Connes system.

The Kronecker–Weber theorem can be seen as a consequence of class field theory, through the properties of the classical Artin homomorphism (3.129).

A modulus $\mathfrak{m}$ is the datum of an integral ideal of the ring of integers $\mathcal{O}_{\mathbb{K}}$ together with a subset of real places. More precisely, it is a product $\mathfrak{m} = \prod_v v^{e_v}$, where $e_v \in \mathbb{N} \cup \{0\}$ is almost everywhere zero, with $e_v \in \{0,1\}$ when v is a real Archimedean place and $e_v = 0$ for v a complex Archimedean place. One associates to a modulus $\mathfrak{m}$ a corresponding ray class group

$$Cl_{\mathfrak{m}}(\mathcal{O}_{\mathbb{K}}) := \mathbb{A}_{\mathbb{K}}^{\mathfrak{m}} / U_{\mathfrak{m}} \mathbb{K}^{\mathfrak{m}}. \tag{3.130}$$

It is obtained as follows: $U_{\mathfrak{m}} \subset \mathbb{A}_{\mathbb{K}}^*$ is of the form $U_{\mathfrak{m}} = \prod_v U_{\mathfrak{m},v}$, where

$$U_{\mathfrak{m},v} = \begin{cases} \mathcal{O}_v^* & e_v = 0 \\ 1 + \mathfrak{m}_v^{e_v} & e_v > 0, \quad \mathfrak{m}_v \subset \mathcal{O}_v \text{ max ideal} \\ \mathbb{R}_+^* & v \text{ real Archimedean place} \end{cases}$$

and we have $\mathbb{A}_{\mathbb{K}}^{\mathfrak{m}} = \{(a_v) \in \mathbb{A}_{\mathbb{K}}^* | \, a_v \in U_{\mathfrak{m},v} \text{ if } e_v > 0\}$ and $\mathbb{K}^m = \mathbb{A}_{\mathbb{K}}^{\mathfrak{m}} \cap \mathbb{K}^*$. For $\mathfrak{m} = 1$, the ray class group is the usual class group $Cl(\mathcal{O}_{\mathbb{K}})$. The ray class field can also be written as a quotient $Cl_{\mathfrak{m}}(\mathcal{O}_{\mathbb{K}}) = I_{\mathbb{K}}^{\mathfrak{m}} / P_{\mathfrak{m}}$. The Artin reciprocity law shows that, given a finite abelian extension $\mathbb{L}$ of a number field $\mathbb{K}$, there exists a modulus $\mathfrak{m}$ such that

$$\mathrm{Ker}(\theta_{\mathbb{L}/\mathbb{K}}) = \mathfrak{n}_{\mathbb{L}/\mathbb{K}}(I_{\mathbb{L}}^{\mathfrak{m}}) P_{\mathfrak{m}}. \tag{3.131}$$

The greatest common divisor of the moduli satisfying this property is called the conductor of $\mathbb{L}/\mathbb{K}$. Moreover, one can prove that if $\mathbb{L}$ is an abelian extension as above, with $\mathfrak{m}$ a modulus satisfying (3.131) and $\mathbb{E}$ is a Galois extension of $\mathbb{K}$ with $\mathfrak{n}_{\mathbb{E}/\mathbb{K}}(I_{\mathbb{E}}^{\mathfrak{m}}) \subset \mathfrak{n}_{\mathbb{L}/\mathbb{K}}(I_{\mathbb{L}}^{\mathfrak{m}}) P_{\mathfrak{m}}$, then $\mathbb{L} \subset \mathbb{E}$. This then gives the Kronecker–Weber theorem: given an abelian extension $\mathbb{L}$ of $\mathbb{Q}$, a modulus $\mathfrak{m}$ satisfying (3.131) can be written as $n \times \infty$, for some $n \in \mathbb{N}$. Take $\mathbb{E} = \mathbb{Q}(\zeta_n)$. This satisfies

$$\mathfrak{n}_{\mathbb{Q}(\zeta_n)/\mathbb{Q}}(I_{\mathbb{Q}(\zeta_n)}^{\mathfrak{m}}) \subset \mathrm{Ker}(\theta_{\mathbb{Q}(\zeta_n)/\mathbb{Q}}) \subset \mathfrak{n}_{\mathbb{L}/\mathbb{Q}}(I_{\mathbb{L}}^{\mathfrak{m}}) P_{\mathfrak{m}} = \mathrm{Ker}(\theta_{\mathbb{L}/\mathbb{Q}}),$$

which by the above implies $\mathbb{L} \subset \mathbb{Q}(\zeta_n)$. For more details see [**6**], [**175**].

In the following treatment of the abelian class field theory of $\mathbb{Q}$ through the Bost–Connes system, we are going to use the identification $\mathbb{Q}^{\mathrm{ab}} = \mathbb{Q}^{\mathrm{cycl}}$. It is an important question to recover a complete proof of the Kronecker–Weber theorem in the context of quantum statistical mechanics.

4.6. KMS states and class field theory.

The main result of Bost–Connes [**30**] is the complete classification of KMS states for the system $(\mathcal{A}_1, \sigma_t)$ and the remarkable intertwining of Galois and symmetries actions on KMS states at zero-temperature. We review these results in this section.

As we have seen in §4.5 above, for a number field $\mathbb{K}$ the *class field theory isomorphism* gives an identification (3.127) of the quotient $C_{\mathbb{K}}/D_{\mathbb{K}}$ of the group of idèle classes $C_{\mathbb{K}} = \mathrm{GL}_1(\mathbb{A}_{\mathbb{K}})/\mathrm{GL}_1(\mathbb{K})$ by the connected component of the identity with the abelianization of the absolute Galois group

$$\mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K}) = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})^{\mathrm{ab}}. \tag{3.132}$$

In the case of $\mathbb{K} = \mathbb{Q}$ the group $C_{\mathbb{Q}}/D_{\mathbb{Q}}$ can be identified with $\mathbb{A}^*_{\mathbb{Q},f}/\mathbb{Q}^*_+ = \hat{\mathbb{Z}}^*$.
As before, we let $\mathcal{E}_\beta$ denote the set of extremal KMS states at inverse temperature β. Then we have the following result.

THEOREM 3.32. ([**30**]) *For the system $(\mathcal{A}_1, \sigma_t)$ described above, the structure of KMS states is the following.*

- *$\mathcal{E}_\beta$ is a singleton for all $0 < \beta \leq 1$. This unique KMS state, restricted to the subalgebra $\mathcal{B}_{\mathbb{Q}}$, takes values*

$$\varphi_\beta(e(a/b)) = f_{-\beta+1}(b)/f_1(b), \qquad \textit{where } f_k(b) = \sum_{d|b} \mu(d)(b/d)^k, \tag{3.133}$$

  with μ the Möbius function, and f_1 is the Euler totient function as in (3.106)*.*
- *For $1 < \beta \leq \infty$, elements of $\mathcal{E}_\beta$ are indexed by the classes of invertible $\mathbb{Q}$-lattices $\rho \in \hat{\mathbb{Z}}^* = \mathrm{GL}_1(\hat{\mathbb{Z}})$, hence by the quotient*

$$\mathcal{E}_\beta \cong \mathrm{GL}_1(\mathbb{Q})\backslash\mathrm{GL}_1(\mathbb{A})/\mathbb{R}^*_+ \cong C_{\mathbb{Q}}/D_{\mathbb{Q}}. \tag{3.134}$$

  Within the range of temperatures $1 < \beta < \infty$, the values of states $\varphi_{\beta,\rho} \in \mathcal{E}_\beta$ on the elements $e(r) \in \mathcal{B}_{\mathbb{Q}}$ are given by polylogarithm functions evaluated at roots of unity, normalized by the Riemann zeta function,

$$\varphi_{\beta,\rho}(e(r)) = \frac{1}{\zeta(\beta)} \mathrm{Li}_\beta(\rho(\zeta_r)), \tag{3.135}$$

  where $\mathrm{Li}_s(z) = \sum_{n=1}^\infty z^n/n^s$.
- *The group $\mathrm{GL}_1(\hat{\mathbb{Z}})$ acts by automorphisms of the system $(\mathcal{A}_1, \sigma_t)$. The induced action of $\mathrm{GL}_1(\hat{\mathbb{Z}})$ on $\mathcal{E}_\beta$ below critical temperature is free and transitive.*
- *The extreme KMS states $\mathcal{E}_\infty$ at zero-temperature have the property that*

$$\varphi(\mathcal{A}_{1,\mathbb{Q}}) \subset \mathbb{Q}^{\mathrm{cycl}}, \quad \forall \varphi \in \mathcal{E}_\beta. \tag{3.136}$$

  Moreover, the class field theory isomorphism (3.127) *intertwines the Galois action on values with the action of $\hat{\mathbb{Z}}^*$ by symmetries,*

$$\gamma\, \varphi(f) = \varphi(\theta^{-1}(\gamma)\, f), \tag{3.137}$$

  for all $\varphi \in \mathcal{E}_\infty$, for all $\gamma \in \mathrm{Gal}(\mathbb{Q}^{\mathrm{cycl}}/\mathbb{Q})$ and for all $f \in \mathcal{A}_{1,\mathbb{Q}}$.

We can write more explicitly the formula for the unique KMS state above critical temperature as

$$\varphi_\beta\,(e(a/b)) = b^{-\beta} \prod_{p\,\mathrm{prime},\, p|b} \left(\frac{1 - p^{\beta-1}}{1 - p^{-1}} \right), \tag{3.138}$$

and for the extremal KMS states below critical temperature as

$$\varphi_{\beta,\rho}(e(a/b)) = \frac{\mathrm{Tr}(\pi_\rho(e(a/b))e^{-\beta H})}{\mathrm{Tr}(e^{-\beta H})} = \frac{1}{\zeta(\beta)} \sum_{n=1}^{\infty} n^{-\beta} \rho(\zeta_{a/b}^n), \tag{3.139}$$

where π_ρ is the representation of the algebra $\mathcal{A}_1$ on the Hilbert space $\mathcal{H} = \ell^2(\mathbb{N})$ given by

$$\pi_\rho(\mu_n)\epsilon_m = \epsilon_{nm}, \quad \pi_\rho(e(a/b))\epsilon_m = \rho(\zeta_{a/b}^m)\epsilon_m, \tag{3.140}$$

where $\rho \in \hat{\mathbb{Z}}^*$ determines an embedding in $\mathbb{C}$ of the cyclotomic field $\mathbb{Q}^{\mathrm{cycl}}$ generated by the abstract roots of unity. The Hamiltonian H implementing the time evolution in this representation is clearly of the form

$$H\epsilon_m = \log(m)\,\epsilon_m, \tag{3.141}$$

so that one sees easily that the partition function is given by the Riemann zeta function,

$$Z(\beta) = \mathrm{Tr}(e^{-\beta H}) = \sum_{n=1}^{\infty} n^{-\beta} = \zeta(\beta). \tag{3.142}$$

This has a pole at $\beta = 1$, where the phase transition occurs.

At zero-temperature, one finds that the values of extremal KMS states on elements of the rational subalgebra $\mathcal{A}_{1,\mathbb{Q}}$ are algebraic numbers. This implies that it is possible to act on these values of KMS states with elements in the Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. What is especially remarkable is the fact that the composition $\gamma \circ \varphi$ obtained in this way, for $\gamma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ extends from a linear functional on $\mathcal{A}_{1,\mathbb{Q}}$ to a *state* on the C^*-algebra $\mathcal{A}_1$. In fact, by the intertwining property (3.137), one knows that the state obtained in this way is the same as the state $\varphi \circ \theta^{-1}(\gamma)$ obtained by the induced action on $\mathcal{E}_\infty$ of the group of symmetries $\hat{\mathbb{Z}}^*$ of the quantum statistical mechanical system $(\mathcal{A}_1, \sigma_t)$. What seems "unreasonable" in this property is the fact that, although elements $\gamma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ extend to automorphisms of $\mathbb{C}$, these are extremely discontinuous and not even Lebesgue measurable (except for $z \mapsto \bar{z}$), and certainly cannot be expected to preserve positivity.

The BC system can be thought of as a quantum statistical mechanical realization of the explicit class field theory of $\mathbb{Q}$. We discuss in §4.7 below how this point of view may help in approaching the more general Hilbert 12th problem of explicit class field theory for other number fields.

4.7. The class field theory problem: algebras and fields.

The presentation of $\mathbb{Q}^{\mathrm{ab}} = \mathbb{Q}^{\mathrm{cycl}}$ is given by the cyclotomic theory, while our rational algebra $\mathcal{A}_{1,\mathbb{Q}}$ is simply the group ring of $\mathbb{Q}/\mathbb{Z}$ with coefficients in $\mathbb{Q}$. We can use this to obtain a presentation of $\mathbb{Q}^{\mathrm{cycl}}$ which does not use any direct reference to field extensions in the following way (cf. [**91**]).

Let $H = \mathbb{Q}/\mathbb{Z}$ be the additive group given by the quotient of $\mathbb{Q}$ by $\mathbb{Z} \subset \mathbb{Q}$. Consider the inclusion of algebras

$$\mathbb{Q}[H] \subset C^*(H) \tag{3.143}$$

and let $J \subset C^*(H)$ be the ideal generated by the idempotents

$$\pi_m = \frac{1}{m} \sum_{r \in H\,,\, m r = 0} u(r), \tag{3.144}$$

where $m > 1$ and the $u(r)$ form the canonical basis of $\mathbb{Q}[H]$.

PROPOSITION 3.33. *With H and J as above, the following holds.*

(1) *The quotient*

$$\mathbb{Q}[H]/(\mathbb{Q}[H] \cap J) \tag{3.145}$$

is a field which is isomorphic to $\mathbb{Q}^{\mathrm{cycl}} \cong \mathbb{Q}^{\mathrm{ab}}$.

(2) *$\mathbb{Q}[H] \cap J$ is equal to the ideal generated by the π_m in $\mathbb{Q}[H]$.*

PROOF. (1) We use the identification of the group $\hat{\mathbb{Z}}$ with the Pontrjagin dual of H given by the pairing (3.53) and the corresponding isomorphism of algebras $C^*(\mathbb{Q}/\mathbb{Z}) \sim C(\hat{\mathbb{Z}})$. It is given explicitly by

$$u(r)(a) = \langle r, a\rangle = \mathbf{e}(ra)\,, \quad \forall a \in \hat{\mathbb{Z}}\,. \tag{3.146}$$

Under this identification, the ideal J corresponds to the ideal of functions that vanish on the closed subset

$$\hat{\mathbb{Z}}^* \subset \hat{\mathbb{Z}}.$$

The quotient map $C^*(H) \to C^*(H)/J$ is just the restriction map

$$f \in C(\hat{\mathbb{Z}}) \mapsto f|_{\hat{\mathbb{Z}}^*}\,. \tag{3.147}$$

Indeed, first notice that the restriction of π_m to $\hat{\mathbb{Z}}^*$ vanishes since the sum of all roots of unity of order m is zero. This shows that J is contained in the kernel of the restriction map (3.147). Moreover, a character of the C^*-algebra $C^*(H)/J$ is given by a point $a \in \hat{\mathbb{Z}}$ on which all elements of J vanish. However, if $a \notin \hat{\mathbb{Z}}^*$, then there exists $m > 1$ such that $a \in m\hat{\mathbb{Z}}$ and one gets $\pi_m(a) = 1 \neq 0$.

The multiplicative group $G = \hat{\mathbb{Z}}^*$ acts by multiplication on both $\mathbb{A}_f/\hat{\mathbb{Z}} \cong H$ and on $\hat{\mathbb{Z}}$ and these actions are compatible with the pairing

$$\langle gr, x\rangle = \langle r, gx\rangle,$$

and therefore with the isomorphism $C(\hat{\mathbb{Z}}) \simeq C^*(\mathbb{Q}/\mathbb{Z})$ and with the actions of G by automorphisms

$$G \subset \mathrm{Aut}\, C^*(H)\,, \quad G \subset \mathrm{Aut}\, C(\hat{\mathbb{Z}}).$$

The group G is also identified with the Galois group of the cyclotomic field, i.e. the subfield $\mathbb{Q}^{\mathrm{cyc}} \subset \mathbb{C}$ generated by all the roots of unity of all orders. For any element $v \in \hat{\mathbb{Z}}^* \subset \hat{\mathbb{Z}}$, let φ_v denote the evaluation map $\varphi_v : C(\hat{\mathbb{Z}}) \to \mathbb{C}$, with $\varphi_v(f) = f(v)$. These maps fulfill the condition

$$g(\varphi_v(f)) = \varphi_v(g(f)) = \varphi_{gv}(f). \tag{3.148}$$

for all $g \in G$ and for all $f \in \mathbb{Q}[H]$, where we use the inclusion (3.143) and the identification $C(\hat{\mathbb{Z}}) \simeq C^*(\mathbb{Q}/\mathbb{Z})$.

Thus, the restriction of $f \in \mathbb{Q}[H]$ to $\hat{\mathbb{Z}}^* \subset \hat{\mathbb{Z}}$ is determined by the single field element $\varphi_1(f) \in \mathbb{Q}^{\mathrm{cycl}}$. The restriction map (3.147) then induces an isomorphism

$$\mathbb{Q}[H]/(\mathbb{Q}[H] \cap J) \xrightarrow{\cong} \mathbb{Q}^{\mathrm{cycl}}\,. \tag{3.149}$$

(2) Let J_0 be the ideal generated by the π_m in $\mathbb{Q}[H]$. Using (3.149) and $J_0 \subset J$ it is enough to show that, for Φ_n the n-th cyclotomic polynomial (cf. [**227**]), one has $\Phi_n(u(1/n)) \in J_0$. When n is prime one has $\Phi_n(u(1/n)) = n\pi_n$. In general, if we let

$$\sigma_k(x) = \sum_{j=0}^{k-1} x^j, \tag{3.150}$$

then the polynomial Φ_n is the g.c.d. of the polynomials $\sigma_m(x^d)$, where $m > 1$ divides n and $d = n/m$. For $x = u(1/n)$, and $n = md$ one has

$$\sigma_m(x^d) = \sum_{j=0}^{m-1} u(j/m) = m\,\pi_m \in J_0$$

Thus, we obtain $\Phi_n(u(1/n)) \in J_0$ as required. □

Corollary 3.34. *Let $f \in C(\hat{\mathbb{Z}}^*)$ be a continuous function. Then for any $\epsilon > 0$ there exists ξ, $\eta \in \mathbb{Q}^{\mathrm{cycl}}$ such that*

$$|f(g) - g(\xi) - ig(\eta)| \leq \epsilon\,, \quad \forall g \in G\,.$$

Proof. The subalgebra $\mathcal{B}_{\mathbb{Q}}(i) = \mathbb{Q}[H] + i\mathbb{Q}[H] \subset C^*(H)$ is norm dense by construction. Thus its image under the restriction map (3.147) is norm dense in $C(\hat{\mathbb{Z}}^*)$ and one can approximate f by an element in the image of $\mathcal{B}_{\mathbb{Q}}(i) = \mathbb{Q}[H] + i\mathbb{Q}[H]$. The image of an element $h \in \mathbb{Q}[H]$ is given by the function

$$g \mapsto g(\xi)\,, \quad \forall g \in G = \hat{\mathbb{Z}}^*$$

where $\xi = \varphi_1(h) \in \mathbb{Q}^{\mathrm{cycl}}$. Thus one gets the required approximation. □

In particular the function of $g \in G$ given by $|g(\xi) + ig(\eta)|$ is largely arbitrary. One can give another direct proof of Corollary 3.34 using for each integer n the density of the field $\mathbb{Q}[\zeta_n]$ of primitive n-th roots of 1 in the product of its completions at the Archimedean places.

All the KMS_∞ states vanish on the ideal J of Proposition 3.33. In fact $\varphi(\pi_n) = \varphi(\mu_n\mu_n^*) = n^{-\beta}$ for any KMS_β state φ. Thus, according to Proposition 3.33, the transition from the algebra $\mathbb{Q}[H]$ to the field $\mathbb{Q}^{\mathrm{cycl}}$ is obtained in the limit of zero-temperature. The extremal KMS_∞ states then give the various embeddings of $\mathbb{Q}^{\mathrm{cycl}}$ in $\mathbb{C}$.

The notion of extremal KMS_∞ state is a generalization of the notion of character of a group and is a fecund source of transcendental functions. It is thus natural to try and extend the above description to that of the maximal abelian extension $\mathbb{K}^{\mathrm{ab}}$ of a number field $\mathbb{K}$.
The first steps are:

- Given a number field $\mathbb{K}$, construct a quantum statistical mechanical system $(\mathcal{A}_{\mathbb{K}}, \sigma_t)$ that has the right partition function (the Dedekind zeta function $\zeta_{\mathbb{K}}(\beta)$) and the right symmetries, which induce an action of $C_{\mathbb{K}}/D_{\mathbb{K}}$ on KMS states.
- Find an arithmetic subalgebra $\mathcal{A}_{\mathbb{K}}^{\mathrm{arithm}}$ which satisfies the intertwining property

$$\varphi \circ \alpha(f) = \theta(\alpha)\varphi(f),$$

between Galois action on values of extremal KMS_∞ states and symmetries, through the class field theory isomorphism θ.

- Compute the presentation of the algebra $\mathcal{A}_{\mathbb{K}}^{\mathrm{arithm}}$ and the extremal KMS_∞ states in terms of their values on the arithmetic subalgebra $\mathcal{A}_{\mathbb{K}}^{\mathrm{arithm}}$.

The first step above has been recently obtained in the desired generality in [**159**], while the second step is still beyond reach at this time. In §8 below we give a complete solution to this problem in the case of imaginary quadratic fields. The next case which is of crucial interest to number theory is the case of real quadratic fields, for which the possibility of an approach via noncommutative geometry was first proposed by Manin in [**216**], [**217**].

4.8. The Shimura variety of $\mathbb{G}_m$.

We give a geometric interpretation of the BC system in terms of the simplest case of a Shimura variety. Shimura varieties constitute a very interesting class of arithmetic varieties that are moduli spaces for Hodge structures and for motives of abelian varieties (cf. e.g. [**112**], [**200**]). Such varieties are obtained as adelic quotients associated to data (G, X) of a reductive group G over $\mathbb{Q}$ and a left homogeneous space X under $G(\mathbb{R})$, in the form

$$Sh(G,X) = G(\mathbb{Q})\backslash(G(\mathbb{A}_{\mathbb{Q},f}) \times X). \tag{3.151}$$

In particular, in this chapter we discuss the simplest cases that correspond to the data $(\mathrm{GL}_1, \{\pm 1\})$ and $(\mathrm{GL}_2, \mathbb{H}^\pm)$, where $\mathbb{H}^\pm = \mathbb{P}^1(\mathbb{C}) \smallsetminus \mathbb{P}^1(\mathbb{R})$. In the first case, as we argue in this section, $Sh(\mathrm{GL}_1, \{\pm 1\})$ is the classical space underlying the BC system, its set of "classical points". In fact, we are going to give an interpretation of the noncommutative space $\mathcal{A}_1$ as a noncommutative version $Sh^{\mathrm{nc}}(\mathrm{GL}_1, \{\pm 1\})$ obtained by adding to the classical space certain noncommutative boundary strata corresponding to degenerating the level structure, in the sense described below. Similarly, we will argue in §7.8 that a similar construction associated to the Shimura variety $Sh(\mathrm{GL}_2, \mathbb{H}^\pm)$ gives a geometric interpretation of the GL_2 generalization of the BC system of [**86**] as a noncommutative $Sh^{\mathrm{nc}}(\mathrm{GL}_2, \mathbb{H}^\pm)$. This point of view was described in [**90**] and generalized to a larger class of Shimura varieties in [**159**], see §9 below.

The Shimura variety associated to the data $(\mathrm{GL}_1, \{\pm 1\})$ is given by the quotient

$$Sh(\mathrm{GL}_1, \{\pm 1\}) = \mathrm{GL}_1(\mathbb{Q})\backslash(\mathrm{GL}_1(\mathbb{A}_f) \times \{\pm 1\}) = \mathbb{A}_f^*/\mathbb{Q}_+^*. \tag{3.152}$$

It is naturally associated to the *cyclotomic tower*, namely the tower of varieties and finite (possibly branched) covering maps which has base $V_1 = \operatorname{Spec}\mathbb{Z}$ and coverings $V_n = \operatorname{Spec}\mathbb{Z}[\zeta_n]$, where ζ_n is a primitive n-th root of unity. The maps are given, for $n|m$, by the ring homomorphisms $\mathbb{Z}[\zeta_n] \hookrightarrow \mathbb{Z}[\zeta_m]$. The group of deck transformations of the cyclotomic tower is the projective limit of the groups $\mathrm{Aut}_{V_1}(V_n)$, which in this case gives

$$\varprojlim_n \mathrm{GL}_1(\mathbb{Z}/n\mathbb{Z}) = \mathrm{GL}_1(\hat{\mathbb{Z}}). \tag{3.153}$$

We have the following set theoretic description of the set of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices up to scaling.

LEMMA 3.35. *The set of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices up to scaling is given by the quotient*

$$\mathrm{GL}_1(\mathbb{Q})\backslash\mathbb{A}^{\cdot}_{\mathbb{Q}}/\mathbb{R}^*_+ = \mathrm{GL}_1(\mathbb{Q})\backslash(\mathbb{A}_{\mathbb{Q},f}\times\{\pm1\}) = \mathbb{A}_{\mathbb{Q},f}/\mathbb{Q}^*_+, \tag{3.154}$$

where $\mathbb{A}^{\cdot}_{\mathbb{Q}} := \mathbb{A}_{\mathbb{Q},f}\times\mathbb{R}^*$ is the set of adèles with nonzero Archimedean component.

PROOF. This follows from the description of the commensurability relation as in Proposition 3.22, in terms of the partially defined action of $\mathbb{Q}^*_+$ on $\hat{\mathbb{Z}}$, where partially defined means that the action $\rho\mapsto r\rho$ is only defined for pairs $(r,\rho)\in\mathcal{U}_1$. Thus, the set of commensurability classes is described set theoretically by the quotient $\hat{\mathbb{Z}}/\mathbb{Q}^*_+$ of this partially defined action. Since $\mathbb{A}_{\mathbb{Q},f} = \hat{\mathbb{Z}}\cdot\mathbb{Q}^*_+$, we can write the quotient $\hat{\mathbb{Z}}/\mathbb{Q}^*_+$ equivalently as the quotient $\mathbb{A}_{\mathbb{Q},f}/\mathbb{Q}^*_+$. □

The quotient (3.154) is not a good quotient in ordinary geometry, but its description as a noncommutative space gives the noncommutative crossed product algebra

$$C_0(\mathbb{A}_f)\rtimes\mathbb{Q}^*_+, \tag{3.155}$$

which, as we have seen in (3.54), is Morita equivalent to the algebra $\mathcal{A}_1$ of the BC system.

The quotient (3.154) with its noncommutative algebra of coordinates (3.155) can then be thought of as the *noncommutative Shimura variety*

$$Sh^{(\mathrm{nc})}(\mathrm{GL}_1,\{\pm1\}) := \mathrm{GL}_1(\mathbb{Q})\backslash(\mathbb{A}_f\times\{\pm1\}) = \mathrm{GL}_1(\mathbb{Q})\backslash\mathbb{A}^{\cdot}/\mathbb{R}^*_+. \tag{3.156}$$

Moreover, the result of [**30**] on the low-temperature KMS states for the BC system shows that the set (3.134) of extremal KMS states at sufficiently low-temperature is identified with the classical Shimura variety

$$\mathcal{E}_\infty = Sh(\mathrm{GL}_1,\{\pm1\}) = \mathbb{A}^*_f/\mathbb{Q}^*_+. \tag{3.157}$$

In the commutative case, evaluating states on a C^*-algebra $C_0(X)$ corresponds to integrating with respect to some probability measure on the underlying space X. In particular, extremal states correspond to measures supported on points. Thus, in the noncommutative setting, when one is dealing with a system $(\mathcal{A},\sigma_t)$ the extremal KMS states provide a good procedure for extracting a notion of "classical points" of a noncommutative space, which here, because of the presence of a nontrivial time evolution, is dependent upon a thermodynamic parameter β. Thus, in the case considered here, one can argue that the extremal KMS states at sufficiently low-temperature identify the classical points $Sh(\mathrm{GL}_1,\{\pm1\})$ of the noncommutative space $Sh^{(\mathrm{nc})}(\mathrm{GL}_1,\{\pm1\})$.

REMARK 3.36. The set of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices can be described as the quotient

$$\mathcal{L}_1 \sim \mathrm{GL}_1(\mathbb{Q})\backslash\mathbb{A}^{\cdot}_{\mathbb{Q}} \sim \mathrm{GL}_1(\mathbb{Q})\backslash(\hat{\mathbb{Z}}\times\mathbb{R}^*) \tag{3.158}$$

where, in the second term, the action of $\mathrm{GL}_1(\mathbb{Q}) = \mathbb{Q}^*$ is only partially defined. One endows $\hat{\mathbb{Z}}\times\mathbb{R}^*$ with the restriction of the additive Haar measure on adèles and lets

$$N = L^\infty(\hat{\mathbb{Z}}\times\mathbb{R}^*)\rtimes\mathrm{GL}_1(\mathbb{Q}) \tag{3.159}$$

be the associated type II_∞ factor. We can use N to obtain a concrete realization of the type II_∞ factor N and the self-adjoint operator Z affiliated to N which were postulated in the construction of the spaces X_z in Chapter 1 §19.2. The semi-finite normal trace on N is given, viewing elements of N as functions on the groupoid $\mathrm{GL}_1(\mathbb{Q}) \ltimes (\hat{\mathbb{Z}} \times \mathbb{R}^*)$ by

$$\mathrm{Tr}_N(f) := \int_{\hat{\mathbb{Z}}\times\mathbb{R}^*} f(1,a)\,da. \tag{3.160}$$

We then define an unbounded element Z of N as the function,

$$Z(1,\rho,\lambda) = \lambda\,, \quad \text{and} \quad Z(r,\rho,\lambda) = 0 \quad \text{if} \quad r \neq 1\,. \tag{3.161}$$

One checks that (with a suitable normalization of da) the spectral measure of Z with respect to the trace Tr_N fulfills (1.905) of Chapter 1 §19.2.

4.9. QSM and QFT of 1-dimensional $\mathbb{Q}$-lattices.

The relation of commensurability on the space of 1-dimensional $\mathbb{Q}$-lattices yields in fact two different noncommutative spaces related by a duality, which corresponds to passing from the quantum statistical mechanical system $(\mathcal{A}_1, \sigma_t)$ to another quantum system obtained by taking the crossed product by the time evolution

$$\hat{\mathcal{A}}_1 = \mathcal{A}_1 \rtimes_\sigma \mathbb{R}. \tag{3.162}$$

We will describe in Chapter IV in a more general setting this duality procedure, which corresponds to the reduction from type III to type II factors introduced by Connes in [**63**]. Here we make some remarks on the properties of these dual spaces in the case of the BC system.

The system $(\mathcal{A}_1, \sigma_t)$ is of quantum statistical mechanical nature, while we argue here that the dual system $\hat{\mathcal{A}}_1$ is in fact quantum field theoretic and it corresponds to the adèle class space that we encountered already in Chapter II in relation to the zeros of the Riemann zeta function.

As we have seen in the previous sections, the first system is obtained by considering the noncommutative algebra of coordinates parameterizing commensurability classes of 1-dimensional $\mathbb{Q}$-lattices up to scaling. On this algebra one has a time evolution σ_t determined by the covolumes of commensurable lattices. Namely, the time evolution has eigenvalues given by the index of pairs of commensurable $\mathbb{Q}$-lattices. The noncommutative space of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices up to scaling is described adelically by the quotient (3.156). The algebra of coordinates, which can be viewed as in [**30**] as the Hecke algebra for a quasi-normal pair of solvable groups, has a regular representation that is of type III_1 (cf. (3.70) above and [**30**]) and determines the time evolution of the system. This time evolution has the set of $\log(p)$, with p a prime number, as set of basic frequencies, as we see from (3.141).

There is a symmetry group acting on this quantum statistical mechanical system, in general by *endomorphisms*, though in the case of the 1-dimensional $\mathbb{Q}$-lattices automorphisms suffice. It is this symmetry that is spontaneously broken at low-temperatures, where the system exhibits several distinct pure phases (extremal KMS states) parameterized by the arithmetic data (3.157). In this 1-dimensional case, the symmetry group is given by the idèle class group modulo the identity component, and we have seen that one has a remarkable intertwining formula between the action of the Galois group $\mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q})$ on the values of KMS_∞ states on arithmetic elements in the algebra and the action of symmetries of the system.

We consider the dual of the GL_1 quantum statistical mechanical system (BC system), under the duality obtained by taking the crossed product by the time evolution. This corresponds to the space $\mathcal{L}_1$ of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices, *not* considered up to scaling. The space $\mathcal{L}_1$ is described geometrically as the total space of a principal $\mathbb{R}_+^*$ bundle, whose base space is the space $\mathcal{L}_1/\mathbb{R}_+^*$ of commensurability classes of $\mathbb{Q}$-lattices up to scaling. The total space $\mathcal{L}_1$ of the fibration has a natural scaling action of $\mathbb{R}_+^*$. This action plays a very important arithmetic role which was described at length in Chapter II. The space $\mathcal{L}_1$ can also be described adelically by the quotient $\mathcal{L}_1 = \mathbb{A}^{\cdot}_{\mathbb{Q}}/\mathbb{Q}^*$.
The algebra $\mathcal{A}_1$ appears as the noncommutative algebra of coordinates of the quotient $\mathcal{L}_1/\mathbb{R}_+^*$ of the space of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices by the commensurability relation. In fact, we have identified $\mathcal{A}_1$ with the groupoid C^*-algebra $C^*(\mathcal{G}_1/\mathbb{R}_+^*)$, cf. Proposition 3.23. The algebra $\hat{\mathcal{A}}_1$ of the dual system (3.162) is associated to the commensurability relation on the space of 1-dimensional $\mathbb{Q}$-lattices, this time not identified up to scaling. Namely, it is the noncommutative algebra of coordinates of the space $\mathcal{L}_1$ and it corresponds to the groupoid $\mathcal{G}_1$.
We have described in Chapter II the spectral realization of the zeros of the Riemann zeta function on the adèle class space $\mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$, which generalizes the simpler model described at the beginning of Chapter II, given by the quantum mechanical system associated to the scaling action of $\mathbb{R}_+^*$ on the phase space of the real line $\mathbb{R}$.
Passing from the space $\mathcal{L}_1 = \mathbb{A}^{\cdot}_{\mathbb{Q}}/\mathbb{Q}^*$ to the full adèle class space $\mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$ amounts to adding the highly singular point $0 \in \mathbb{R}$ at the Archimedean component, which plays an important role in terms of its contribution to the trace formula of [**71**]. The adèle class space $\mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$ should be thought of as the configuration space for a quantum field theory, where the infinitely many degrees of freedom correspond to the places of $\mathbb{Q}$, i.e. the prime numbers and the Archimedean place at infinity. The field configurations are given by the adèles, with the space $\mathbb{A}_{\mathbb{Q}}$ divided by the action of the gauge group $\mathrm{GL}_1(\mathbb{Q})$.
When one considers, as we discussed in Chapter II, only finitely many degrees of freedom (the semi-local case) and in particular only the place at infinity, then we have seen that the semiclassical approximation exhibits the main terms in the asymptotic formula for the number of zeros of the Riemann zeta function. The Lefschetz formula for the scaling action recovers the Riemann–Weil explicit formula as a semiclassical approximation. The difficulty in extending this calculation to the global case lies in the quantum field theoretic problem of passing to infinitely many degrees of freedom.

As we have seen, the zeros of zeta appear as an absorption spectrum, namely as lacunae in a continuous spectrum, where the width of the absorption lines depends on the presence of a cutoff. The full idèle class group appears as symmetries of the system and L-functions with Grössencharakter replace the Riemann zeta function in nontrivial sectors. The $\log(p)$ appear as periods of the orbits of the scaling action on the quotient $\mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$.
The properties of these dual systems are summarized in Table 1.

5. 2-dimensional $\mathbb{Q}$-lattices

We now describe an analog of the BC system based on commensurability classes of 2-dimensional $\mathbb{Q}$-lattices. This system was introduced in [**86**].

Quantum statistical mechanics	Quantum field theory
Commensurability classes of $\mathbb{Q}$-lattices modulo scaling	Commensurability classes of $\mathbb{Q}$-lattices
$\mathcal{A} = C^*(\mathbb{Q}/\mathbb{Z}) \rtimes \mathbb{N}^\times$	$\mathcal{A} \rtimes_{\sigma_t} \mathbb{R}$
Time evolution σ_t	Energy scaling $U(\lambda)$, $\lambda \in \mathbb{R}_+^*$
$\{\log p\}$ as frequencies	$\{\log p\}$ as periods of orbits
Arithmetic rescaling μ_n	Renormalization group flow $\mu\partial_\mu$
Symmetry group $\hat{\mathbb{Z}}^*$ as Galois action on $T = 0$ states	Idèle class group as gauge group
System at zero-temperature	$\mathrm{GL}_1(\mathbb{Q})\backslash\mathrm{GL}_1(\mathbb{A}_\mathbb{Q})$
System at critical temperature (Riemann's ζ as partition function)	Spectral realization (Zeros of ζ as absorption spectrum)
Type III_1	Type II_∞

TABLE 1. Dual systems

PROPOSITION 3.37. *The set of 2-dimensional $\mathbb{Q}$-lattices is identified with the quotient*

$$\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R})), \tag{3.163}$$

where $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. The set of 2-dimensional $\mathbb{Q}$-lattices up to scaling is identified with the quotient

$$\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H}), \tag{3.164}$$

with $\mathbb{H}$ the upper half plane.
Given $\rho \in M_2(\hat{\mathbb{Z}})$ and $z \in \mathbb{H}$ the $\mathbb{Q}$-lattice up to scale associated to the pair (ρ, z) is given by

$$(\Lambda, \phi) = (\lambda(\mathbb{Z} + \mathbb{Z}z), \lambda(\rho_1 - z\rho_2)), \tag{3.165}$$

for any $\lambda \in \mathbb{C}^*$.

PROOF. We use the basis $\{e_1 = 1, e_2 = -i\}$ of $\mathbb{C}$ as a 2-dimensional $\mathbb{R}$-vector space to let $\mathrm{GL}_2^+(\mathbb{R})$ act on $\mathbb{C}$ as $\mathbb{R}$-linear transformations. Thus, for

$$\alpha(xe_1 + ye_2) = (ax+by)e_1 + (cx+dy)e_2, \quad \text{with } \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R}), \tag{3.166}$$

we define

$$\Lambda_0 := \mathbb{Z}e_1 + \mathbb{Z}e_2 = \mathbb{Z} + i\mathbb{Z}. \tag{3.167}$$

We have an identification $M_2(\hat{\mathbb{Z}}) = \mathrm{Hom}(\mathbb{Q}^2/\mathbb{Z}^2, \mathbb{Q}^2/\mathbb{Z}^2)$. Equivalently, with our choice of basis, we view an element $\rho \in M_2(\hat{\mathbb{Z}})$ as a map

$$\rho : \mathbb{Q}^2/\mathbb{Z}^2 \to \mathbb{Q}\Lambda_0/\Lambda_0, \qquad \rho(a) = \rho_1(a)e_1 + \rho_2(a)e_2\,. \tag{3.168}$$

Every 2-dimensional $\mathbb{Q}$-lattice (Λ, ϕ) can then be described by data

$$(\Lambda, \phi) = (\alpha^{-1}\Lambda_0, \alpha^{-1}\rho), \tag{3.169}$$

for some $\alpha \in \mathrm{GL}_2^+(\mathbb{R})$ and some $\rho \in M_2(\hat{\mathbb{Z}})$. The equality

$$(\alpha^{-1}\Lambda_0, \alpha^{-1}\rho) = (\beta^{-1}\Lambda_0, \beta^{-1}\sigma)$$

holds if and only if $\beta\alpha^{-1}\Lambda_0 = \Lambda_0$ and $\beta\alpha^{-1}\rho = \sigma$. The first equality means that $\gamma = \beta\alpha^{-1} \in \mathrm{SL}_2(\mathbb{Z})$ and the second that $\sigma = \gamma\rho$. Thus $(\beta, \sigma) = (\gamma\alpha, \gamma\rho)$ and the space of 2-dimensional $\mathbb{Q}$-lattices can be described as the quotient (3.163) by the diagonal action of $\Gamma = \mathrm{SL}_2(\mathbb{Z})$.
Under the identification (3.166) the action of $\lambda \in \mathbb{C}^*$ by multiplication corresponds to the matrix:

$$\lambda = a + ib \in \mathbb{C}^* \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R}) \tag{3.170}$$

and we identify $\mathbb{C}^*$ as a subgroup of $\mathrm{GL}_2^+(\mathbb{R})$ via this map. The action of $\mathbb{C}^*$ by scaling on $\mathbb{Q}$-lattices is given by

$$\lambda(\Lambda, \phi) = (\lambda\Lambda, \lambda\phi)\,, \quad \forall \lambda \in \mathbb{C}^* \tag{3.171}$$

and this corresponds to the right multiplication of α by λ^{-1} in (3.169), while ρ is left unchanged. Thus the set of 2-dimensional $\mathbb{Q}$-lattices up to scaling is identified with the quotient

$$\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R})/\mathbb{C}^*) = \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H}). \tag{3.172}$$

More specifically, the group $\mathrm{GL}_2^+(\mathbb{R})$ acts on the upper half plane $\mathbb{H}$ by fractional transformations,

$$\alpha(z) = \frac{az+b}{cz+d}\,, \quad \forall \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R}) \tag{3.173}$$

We identify the quotient $\mathrm{GL}_2^+(\mathbb{R})/\mathbb{C}^*$ with $\mathbb{H}$ via the map

$$\alpha \in \mathrm{GL}_2^+(\mathbb{R}) \mapsto z = \alpha(i) \in \mathbb{H}. \tag{3.174}$$

This gives the required identification (3.164).
Let us check (3.165). The right $\mathbb{C}^*$-coset associated to $z = x + iy \in \mathbb{H}$ contains the matrix

$$\alpha = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R}).$$

One has

$$\alpha^{-1} = y^{-1} \begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix},$$

which one can replace up to scale by

$$g = \begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix}.$$

One has $g(e_1) = e_1 = 1$ and $g(e_2) = -xe_1 + ye_2 = -z$ which gives the required answer. $\square$

The commensurability relation on the space of 2-dimensional $\mathbb{Q}$-lattices is implemented by a partially defined action of $\mathrm{GL}_2^+(\mathbb{Q})$. Namely, for $g \in \mathrm{GL}_2^+(\mathbb{Q})$ and $(\Lambda, \phi) = (\alpha^{-1}\Lambda_0, \alpha^{-1}\rho)$ such that $g\rho \in M_2(\hat{\mathbb{Z}})$, the $\mathbb{Q}$-lattice $(\alpha^{-1}g^{-1}\Lambda_0, \alpha^{-1}\rho)$ is commensurable to (Λ, ϕ). Moreover, one can see as in the proof of Proposition 3.22 that all $\mathbb{Q}$-lattices commensurable to a given (Λ, ϕ) are of this form. Thus, at the set theoretic level, the space of commensurability classes of 2-dimensional $\mathbb{Q}$-lattices is the quotient of the space $M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R})$ by the partially defined action of $\mathrm{GL}_2^+(\mathbb{Q})$ given by

$$g(\rho, \alpha) = (g\rho, g\alpha) \quad \textit{if} \quad g\rho \in M_2(\hat{\mathbb{Z}})\,. \tag{3.175}$$

Note that the equivalence relation $(\rho, \alpha) \sim (g\rho, g\alpha)$ generated by this partially defined action of $\mathrm{GL}_2^+(\mathbb{Q})$ is finer than the equivalence relation for the orbits of Γ. In this way the space of commensurability classes of 2-dimensional $\mathbb{Q}$-lattices is a quotient of the space

$$\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R}))$$

of (3.163). We are going to describe this more precisely in terms of groupoids and noncommutative spaces in §5.2 below.

5.1. Elliptic curves and Tate modules.

We first give a geometric reinterpretation of the data of 2-dimensional $\mathbb{Q}$-lattices and of commensurability in terms of elliptic curves and Tate modules, following [**91**].

For an elliptic curve $E = \mathbb{C}/\Lambda$ let $E[n] = (\frac{1}{n}\Lambda)/\Lambda \sim \Lambda/n\Lambda$ be the n-torsion points of E. The ℓ-adic Tate module of E is defined as

$$T_\ell E = \varprojlim_n E[\ell^n] = \Lambda \otimes \mathbb{Z}_\ell. \tag{3.176}$$

One can also define the total (or adelic) Tate module of E as

$$TE = \varprojlim_n E[n] = \Lambda \otimes \hat{\mathbb{Z}}. \tag{3.177}$$

In terms of the torsion points $E[n] = (\frac{1}{n}\Lambda)/\Lambda$ the map from $E[n]$ to $E[m]$ for $n = md$ is given by $a \mapsto da$. Thus one has an equivalent description in terms of the torsion subgroup E_{tor} of the elliptic curve:

$$TE = \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, E_{\mathrm{tor}})\,. \tag{3.178}$$

Given $\rho \in \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, E_{\mathrm{tor}})$ the corresponding element of TE is given by the $(\rho(\frac{1}{n}))_{n\in\mathbb{N}}$.

PROPOSITION 3.38. *The data of a 2-dimensional $\mathbb{Q}$-lattice up to scaling are equivalent to the data of an elliptic curve E, together with a pair of points $\xi = (\xi_1, \xi_2)$ in the total Tate module TE. Commensurability of 2-dimensional $\mathbb{Q}$-lattices up to scale is then implemented by an isogeny of the corresponding elliptic curves, with the elements ξ and ξ' related via the induced map on the corresponding Tate modules.*

PROOF. To a 2-dimensional $\mathbb{Q}$-lattice (Λ, ϕ) we first associate the elliptic curve $E = \mathbb{C}/\Lambda$. The subgroup $\mathbb{Q}\Lambda/\Lambda$ of $E = \mathbb{C}/\Lambda$ is the torsion subgroup E_{tor} of the elliptic curve. Thus, one can rewrite the map ϕ as a map $\mathbb{Q}^2/\mathbb{Z}^2 \to E_{\text{tor}}$. By taking $\text{Hom}(\mathbb{Q}/\mathbb{Z}, -)$, we rewrite the map ϕ as a $\hat{\mathbb{Z}}$-linear map

$$\hat{\mathbb{Z}} \oplus \hat{\mathbb{Z}} \to TE, \tag{3.179}$$

that is, a choice of two elements (ξ_1, ξ_2) of the latter. Commensurability of $\mathbb{Q}$-lattices up to scale is rephrased as the condition that the elliptic curves are isogenous and the points in the Tate module are related via the induced map. $\square$

Thus we have obtained the equivalent description:

$$\begin{gathered}\text{Commensurability classes of } \mathbb{Q}\text{-lattices up to scale}\\ \Longleftrightarrow \text{Isogeny classes of data } (E, \xi_1, \xi_2).\end{gathered}$$

This corresponds to the following set theoretic description of the commensurability classes of 2-dimensional $\mathbb{Q}$-lattices up to scaling.

PROPOSITION 3.39. *The space of commensurability classes of 2-dimensional $\mathbb{Q}$-lattices up to scaling is given by the quotient*

$$\text{GL}_2^+(\mathbb{Q})\backslash(M_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}). \tag{3.180}$$

for the diagonal action of $\text{GL}_2^+(\mathbb{Q})$ *given by* $(\rho, z) \mapsto (g\rho, g(z))$.

PROOF. We already saw that the space $\mathcal{L}_2$ of commensurability classes of 2-dimensional $\mathbb{Q}$-lattices is the quotient of $M_2(\hat{\mathbb{Z}}) \times \text{GL}_2^+(\mathbb{R})$ by the partially defined action of $\text{GL}_2^+(\mathbb{Q})$ given by (3.175). The equality

$$M_2(\mathbb{A}_{\mathbb{Q},f}) = \text{GL}_2^+(\mathbb{Q}) M_2(\hat{\mathbb{Z}})$$

shows that $\mathcal{L}_2$ can equivalently be described as the quotient

$$\text{GL}_2^+(\mathbb{Q})\backslash(M_2(\mathbb{A}_{\mathbb{Q},f}) \times \text{GL}_2^+(\mathbb{R})). \tag{3.181}$$

Modding out by the scaling action of $\mathbb{C}^*$ one obtains (3.180). $\square$

One can also prove Proposition 3.39 more geometrically using Proposition 3.38. Indeed, let us identify $(\hat{\mathbb{Z}} \oplus \hat{\mathbb{Z}}) \otimes \mathbb{Q}$ with $\mathbb{A}_{\mathbb{Q},f} \oplus \mathbb{A}_{\mathbb{Q},f}$ and $TE \otimes \mathbb{Q} = (\Lambda \otimes \hat{\mathbb{Z}}) \otimes \mathbb{Q} = \Lambda \otimes \mathbb{A}_{\mathbb{Q},f}$. The data of Proposition 3.38 can be rephrased as isogeny classes of data (E, η) of an elliptic curve together with an $\mathbb{A}_{\mathbb{Q},f}$-homomorphism

$$\eta : \mathbb{A}_{\mathbb{Q},f} \oplus \mathbb{A}_{\mathbb{Q},f} \to \Lambda \otimes \mathbb{A}_{\mathbb{Q},f}, \tag{3.182}$$

The commensurability of $\mathbb{Q}$-lattices corresponds to the equivalence $(E, \eta) \sim (E', \eta')$ given by an isogeny $g : E \to E'$ with $\eta' = (g \otimes 1) \circ \eta$. Thus, the equivalence classes are identified with the quotient of $M_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}^{\pm}$ by the action of $\text{GL}_2(\mathbb{Q})$ given by $(\rho, z) \mapsto (g\rho, g(z))$.

5.2. Algebras and groupoids.

We let $\mathcal{G}_2$ denote the groupoidof the equivalence relation of commensurability on the space of 2-dimensional $\mathbb{Q}$-lattices. We have the following description, which is analogous to the 1-dimensional case.
Consider first the groupoid$\tilde{\mathcal{G}}_2$ defined as the set

$$\tilde{\mathcal{G}}_2 = \{(g,\rho,\alpha) \in \mathrm{GL}_2^+(\mathbb{Q}) \times M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R}) \,|\, g\rho \in M_2(\hat{\mathbb{Z}})\}, \tag{3.183}$$

with space of units $\tilde{\mathcal{G}}_2^{(0)} = M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R})$, source and range maps $s(g,\rho,\alpha) = (\rho,\alpha)$ and $r(g,\rho,\alpha) = (g\rho, g\alpha)$. The composition is given by

$$(g_1,\rho_1,\alpha_1) \circ (g_2,\rho_2,\alpha_2) = (g_1g_2,\rho_2,\alpha_2), \quad \text{if } (g_2\rho_2, g_2\alpha_2) = (\rho_1,\alpha_1). \tag{3.184}$$

For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is an action of $\Gamma \times \Gamma$ on (3.183) of the form

$$(g,\rho,\alpha) \mapsto (\gamma_1 g \gamma_2^{-1}, \gamma_2\rho, \gamma_2\alpha). \tag{3.185}$$

We consider then the quotient of the groupoid (3.183) by the action (3.185) of $\Gamma\times\Gamma$.

$$(\Gamma \times \Gamma)\backslash\tilde{\mathcal{G}}_2 \sim \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q}) \times_\Gamma (M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2(\mathbb{R}))\,. \tag{3.186}$$

The space of units is $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R}))$ and one checks that the above source and range maps pass to the quotient. In fact $s(\gamma_1 g\gamma_2^{-1}, \gamma_2\rho, \gamma_2\alpha) = \gamma_2 s(g,\rho,\alpha)$ and $r(\gamma_1 g\gamma_2^{-1}, \gamma_2\rho, \gamma_2\alpha) = \gamma_1\, r(g,\rho,\alpha)$. Given a composable pair (δ,δ') of elements in $(\Gamma\times\Gamma)\backslash\tilde{\mathcal{G}}_2$, i.e. a pair such that $s(\delta) = r(\delta')$, one can lift it to a composable pair in $\tilde{\mathcal{G}}_2$. Indeed, let δ be the class of the $(\gamma_1 g\gamma_2^{-1}, \gamma_2\rho, \gamma_2\alpha)$ and similarly for δ'. Then since (ρ,α) and $(g'\rho', g'\alpha')$ have the same class in $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2^+(\mathbb{R}))$, one can choose γ_2 and γ_1' such that

$$s(\gamma_1 g\gamma_2^{-1}, \gamma_2\rho, \gamma_2\alpha) = r(\gamma_1' g' \gamma_2'^{-1}, \gamma_2'\rho', \gamma_2'\alpha')\,.$$

Moreover the ensuing equality $\gamma_2\alpha = \gamma_1' g'\alpha'$ shows that the only freedom left is to multiply both γ_2 and γ_1' on the left by the same element $\gamma \in \Gamma$. It follows that the class of the composition of the lifted pair is well defined. It is given by

$$\delta\circ\delta' = (\gamma_1 g\gamma_2^{-1}\gamma_1' g'\gamma_2'^{-1}, \gamma_2'\rho', \gamma_2'\alpha')\,.$$

This suffices to show that this composition law turns $(\Gamma\times\Gamma)\backslash\tilde{\mathcal{G}}_2$ into a groupoid. It is locally compact and etale by construction.

Proposition 3.40. *The following is an isomorphism of locally compact étale groupoids between* $(\Gamma\times\Gamma)\backslash\tilde{\mathcal{G}}_2$ *and* $\mathcal{G}_2$

$$\tilde{\gamma}(g,\rho,\alpha) = ((\alpha^{-1}g^{-1}\Lambda_0, \alpha^{-1}\rho), (\alpha^{-1}\Lambda_0, \alpha^{-1}\rho)). \tag{3.187}$$

for Λ_0 *as in* (3.167).

Proof. The argument is analogous to the proof of Proposition 3.22, using the result of Proposition 3.37. ☐

One can consider the associated convolution algebra $C^*(\mathcal{G}_2)$. Notice that in this case, this is not a unital algebra since the space of units is not compact. A function on $\mathcal{G}_2$ can be viewed as a $\Gamma \times \Gamma$-invariant function on $\tilde{\mathcal{G}}_2$. The convolution product is then of the form

$$f_1 * f_2\,(g,\rho,\alpha) = \sum_{h\in\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})\,:\,h\rho\in M_2(\hat{\mathbb{Z}})} f_1(gh^{-1}, h\rho, h\alpha) f_2(h,\rho,\alpha), \tag{3.188}$$

where the pair $((gh^{-1}, h\rho, h\alpha),(h,\rho,\alpha))$ is composable in $\tilde{\mathcal{G}}_2$ by construction. Note that the freedom of left multiplication of h by an element of Γ given by the choice of the lift from $\mathcal{G}_2$ to $\tilde{\mathcal{G}}_2$ is taken care of by the summation on the coset space $\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$. This is possible because of the $\Gamma\times\Gamma$-invariance of the functions f_j. One checks that $f_1 * f_2$ is also $\Gamma \times \Gamma$-invariant.

Things differ from the 1-dimensional case when we consider the effect on $\mathcal{G}_2$ of the scaling action, namely we consider the quotient

$$\mathcal{U}_2 := \mathcal{G}_2/\mathbb{C}^*. \tag{3.189}$$

In fact, the scaling action of $\mathbb{C}^*$ does preserve the commensurability relation, but in the resulting quotient (3.189) it is no longer possible to get a groupoid law from the groupoid law in $\mathcal{G}_2$. The reason is the presence of lattices with nontrivial automorphisms, so that the scaling action of $\mathbb{C}^*$ is not free on the units $\mathcal{G}_2^{(0)}$. This creates an ambiguity in the composition law as follows:

Remark 3.41. *The composition law in the quotient $\mathcal{G}_2/\mathbb{C}^*$ is not well defined: composable lifts of a composable pair (δ,δ') can yield inequivalent results.*

Proof. This can be seen by looking at the following simple case. Consider the $\mathbb{Q}$-lattices given by

$$(\Lambda,\phi) = (\mathbb{Z}+2i\mathbb{Z}, 0), \qquad (\Lambda',\phi') = (\mathbb{Z}+i\mathbb{Z},0). \tag{3.190}$$

They are commensurable. The composition

$$((\Lambda,\phi),(\Lambda',\phi'))\circ((\Lambda',\phi'),(\Lambda,\phi))$$

in the groupoid$\mathcal{G}_2$ is the unit $((\Lambda,\phi),(\Lambda,\phi))$. In the groupoid $\mathcal{G}_2$ we can also consider the composition

$$(\lambda(\Lambda,\phi),\lambda(\Lambda',\phi'))\circ((\Lambda',\phi'),(\Lambda,\phi)),$$

for $\lambda = i \in \mathbb{C}^*$, since for the data (3.190) we have $\phi = \phi' = 0$ and $i\Lambda' = \Lambda'$, so that $i(\Lambda',\phi') = (\Lambda',\phi')$. This composition is not a unit in $\mathcal{G}_2$, since $i(\Lambda,\phi) \neq (\Lambda,\phi)$. Moreover even after division by $\mathbb{C}^*$ one gets that $((\Lambda,\phi),(\Lambda,\phi))$ is not equivalent to $(i(\Lambda,\phi),(\Lambda,\phi))$. But these correspond to compositions of composable lifts of the pair (δ,δ') given by the class modulo $\mathbb{C}^*$ of

$$\delta = ((\Lambda,\phi),(\Lambda',\phi'))\,, \quad \delta' = ((\Lambda',\phi'),(\Lambda,\phi))$$

∎

Thus, the groupoid structure does not pass to the quotient $\mathcal{G}_2/\mathbb{C}^*$, because of the presence of lattices (like the one corresponding to the point $z = i \in \mathbb{H}$) that have nontrivial automorphisms.
However, even though $\mathcal{G}_2/\mathbb{C}^*$ is not a groupoid, we can still associate to this quotient by scaling a convolution algebra obtained from the groupoid algebra of $\mathcal{G}_2$, using the

convolution product (3.188) on functions $f(g,\rho,\alpha)$ that are invariant under the $\Gamma\times\Gamma$ action (3.185). The convolution product

$$f_1 * f_2\,(g,\rho,\alpha) = \sum_{s\in\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})\,:\,s\rho\in M_2(\hat{\mathbb{Z}})} f_1(gs^{-1},s\rho,s\alpha)f_2(s,\rho,\alpha), \tag{3.191}$$

and the involution which is of the form $f^*(g,\rho,\alpha)=\overline{f(g^{-1},g\rho,g\alpha)}$ will be applied to functions which are no longer of compact support in the groupoid $\mathcal{G}_2$ but which are invariant under the scaling action of $\mathbb{C}^*$. As in the 1-dimensional case, we say that a function $f(g,\rho,\alpha)$ is homogeneous of weight k with respect to the scaling action of $\mathbb{C}^*$ if it has the property that

$$f(g,\rho,\alpha\lambda)=\lambda^k f(g,\rho,\alpha), \quad \forall\lambda\in\mathbb{C}^*, \tag{3.192}$$

where on the left hand side $\lambda\in\mathbb{C}^*$ is regarded as an element in $\mathrm{GL}_2(\mathbb{R})$ via the map (3.170).

We then define the convolution algebra of the quotient $\mathcal{U}_2=\mathcal{G}_2/\mathbb{C}^*$ by restricting the convolution product (3.191) to weight zero functions on $\mathcal{G}_2$. The associativity of the product follows as long as we check that the sums involved in (3.191) are finite. To ensure that this is the case we need to define a suitable notion of function with compact support on $\mathcal{U}_2$. Note that the action

$$(g,\rho,z)\mapsto(\gamma_1 g\gamma_2^{-1},\gamma_2\rho,\gamma_2(z)), \tag{3.193}$$

of $\Gamma\times\Gamma$ on

$$\tilde{\mathcal{U}_2}=\{(g,\rho,z)\in \mathrm{GL}_2^+(\mathbb{Q})\times M_2(\hat{\mathbb{Z}})\times\mathbb{H}\,:\, g\rho\in M_2(\hat{\mathbb{Z}})\} \tag{3.194}$$

is still proper since the action of $\Gamma=\mathrm{SL}_2(\mathbb{Z})$ on the upper half plane $\mathbb{H}$ is the usual one by fractional linear transformation and is proper. Thus the quotient

$$\mathcal{U}_2=(\Gamma\times\Gamma)\backslash\tilde{\mathcal{U}_2} \tag{3.195}$$

is a locally compact space.

We let $\mathcal{A}_c(\mathcal{U}_2)$ be the algebra of functions with compact support on $\mathcal{U}_2$ with the convolution product given as above by

$$f_1 * f_2\,(g,\rho,z) = \sum_{s\in\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})\,:\,s\rho\in M_2(\hat{\mathbb{Z}})} f_1(gs^{-1},s\rho,s(z))f_2(s,\rho,z), \tag{3.196}$$

and the involution by $f^*(g,\rho,z)=\overline{f(g^{-1},g\rho,g(z))}$.

By construction functions $f\in\mathcal{A}_c(\mathcal{U}_2)$ are functions $f(g,\rho,z)$ on $\tilde{\mathcal{U}_2}$ satisfying

$$f(\gamma g,\rho,z)=f(g,\rho,z), \qquad f(g\gamma,\rho,z)=f(g,\gamma\rho,\gamma(z)), \tag{3.197}$$

for all $\gamma\in\Gamma$ and all $(g,\rho,z)\in\tilde{\mathcal{U}_2}$. We can encode f by the functions f_g on $M_2(\hat{\mathbb{Z}})\times\mathbb{H}$ defined by $f_g(\rho,z)=f(g,\rho,z)$. By (3.197) f_g only depends on the left coset of $g\in\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$. Moreover

$$f_{g\gamma}(\rho,z)=f_g(\gamma\rho,\gamma(z))\,. \tag{3.198}$$

This shows that f is uniquely determined by the f_g on a set of representatives for the double coset space $\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})/\Gamma$. Moreover the compactness of the support of $f\in\mathcal{A}_c(\mathcal{U}_2)$ implies that the support of f is finite in $\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})/\Gamma$ and since the inclusion $\Gamma\subset\mathrm{GL}_2^+(\mathbb{Q})$ fulfills the commensurability condition (3.62) it follows that the support of f is finite in $\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$. This shows that the sum involved in (3.196) is finite and also allows one to check that the convolution product belongs to $\mathcal{A}_c(\mathcal{U}_2)$. In fact, let us introduce the following notation for functions f on $M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$:

(3.199)
$$(f|s)(\rho, z) = f(s\rho, s(z)) \text{ for } s\rho \in M_2(\hat{\mathbb{Z}}) \text{ and } (f|s)(\rho, z) = 0 \text{ for } s\rho \notin M_2(\hat{\mathbb{Z}})\,.$$

We can then write the convolution product in terms of the functions f_g as

$$(f * f')_g = \sum_{s \in \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})} f_{gs^{-1}}|s\; f'_s \tag{3.200}$$

The map $g \mapsto f_g$ has finite support in $\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$ and the functions f_g themselves are invariant under a congruence subgroup.

LEMMA 3.42. *Elements $f \in \mathcal{A}_c(\mathcal{U}_2)$ have the property that the function f_g defined by $f_g(\rho, z) = f(g, \rho, z)$ depends on (ρ, z) only through the orbit of (ρ, z) under the congruence subgroup $\Gamma \cap g^{-1}\Gamma g$.*

PROOF. One has for $\gamma \in \Gamma$ and $\gamma' = g\gamma g^{-1} \in \Gamma$,

$$\begin{aligned} f_g(\gamma\rho, \gamma(z)) &= f(g, \gamma\rho, \gamma(z)) = f(g\gamma, \rho, z) \\ &= f(\gamma' g, \rho, z) = f(g, \rho, z) = f_g(\rho, z)\,. \end{aligned}$$

□

We now define representations π_y of $\mathcal{A}_c(\mathcal{U}_2)$ associated to elements $y \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ as follows. We let, for $y = (\rho, z)$,

$$G_y := \{g \in \mathrm{GL}_2^+(\mathbb{Q}) \,:\, g\rho \in M_2(\hat{\mathbb{Z}})\}. \tag{3.201}$$

We note that in (3.196) the value of $f_1 * f_2\,(g, \rho, z)$ depends on f_2 only through its restriction to the $\{(s, \rho, z)\,,\ s \in G_y\}$. Thus the following defines a representation π_y of $\mathcal{A}_c(\mathcal{U}_2)$ on the Hilbert space $\ell^2(\Gamma\backslash G_y)$,

$$(\pi_y(f)\xi)(g) := \sum_{h \in \Gamma\backslash G_y} f(gh^{-1}, h\rho, h(z))\xi(h), \qquad \forall g \in G_y, \tag{3.202}$$

for all left Γ-invariant functions ξ on G_y.

DEFINITION 3.43. *The C^*-algebra $\mathcal{A}_2$ is obtained as the norm completion of the convolution algebra $\mathcal{A}_c(\mathcal{U}_2)$ with respect to the norm*

$$\|f\| := \sup_y \|\pi_y(f)\|. \tag{3.203}$$

The compactness of the support of f ensures that the supremum in (3.203) is finite. Since $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ is only locally compact but not compact, the function

$$f(g, \rho, z) = 1 \text{ for } g \in \Gamma \text{ and } f(g, \rho, z) = 0 \text{ for } g \notin \Gamma,$$

which would be the unit, does not have compact support and the algebra $\mathcal{A}_2$ is non-unital.

The construction of the representations π_y of the algebra $\mathcal{A}_c(\mathcal{U}_2)$ corresponds to the general construction described in §4 of the representation of the convolution algebra of the etale groupoid associated to an equivalence relation with countable orbits. To each equivalence class c one associates the natural action by left convolution in the Hilbert space $\ell^2(c)$. In our case the equivalence relation is that of commensurability on the space of $\mathbb{Q}$-lattices up to scale, which we identify with $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ by Proposition 3.37 and (3.165). We let p be the quotient map

$$p_\Gamma : M_2(\hat{\mathbb{Z}}) \times \mathbb{H} \to \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H}).$$

For $x \in \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ let $c(x)$ denote the commensurability class of the $\mathbb{Q}$-lattice up to scale determined by x.

LEMMA 3.44. *Let $y = (\rho, z) \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ be an element in the preimage of x under p. The map*

$$G_y \ni g \mapsto p_\Gamma(gy) \in \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$$

is a surjection from $\Gamma\backslash G_y$ to $c(x)$.

PROOF. Indeed the commensurability classes are the orbits under the partial action of $\mathrm{GL}_2^+(\mathbb{Q})$. $\square$

In general the above map need not be an injection. For instance injectivity fails in the case of an element $y = (0, z)$ where $z \in \mathbb{H}$ is a complex multiplication point. This can be seen by the following result.

LEMMA 3.45. *Suppose given $g \in \mathrm{GL}_2^+(\mathbb{Q})$, $g \neq 1$. Then any element $y = (\rho, z) \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ satisfying $gy = y$ has $\rho = 0$. Moreover, the element $g \in \mathrm{GL}_2^+(\mathbb{Q})$ satisfying $gy = y$ is contained in the subgroup $\mathbb{Q}^* \subset \mathrm{GL}_2^+(\mathbb{Q})$, unless $z \in \mathbb{H}$ is a complex multiplication point, in which case $g \in \mathbb{K}^*$, where $\mathbb{K}$ is the imaginary quadratic field $\mathbb{K} = \mathbb{Q}(z)$ and $\mathbb{K}^*$ is embedded in $\mathrm{GL}_2^+(\mathbb{Q})$ by the choice of the basis $\{z, 1\}$ of K over $\mathbb{Q}$.*

PROOF. To prove the first statement it is enough to show that given $g \in \mathrm{GL}_2^+(\mathbb{Q})$, $g \neq 1$ with $g(z) = z$ for some $z \in \mathbb{H}$ one has $\det(g-1) \neq 0$. Then one has $g - 1 \in \mathrm{GL}_2(\mathbb{Q})$ and $g\rho = \rho$ can be rewritten as $(g-1)\rho = 0$ which implies $\rho = 0$. We write the element $g \in \mathrm{GL}_2^+(\mathbb{Q})$ as

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{3.204}$$

The equation $g(z) = z$ means $az + b = z(cz + d)$. If $c \neq 0$ this implies that z is an imaginary quadratic number, a solution of $P(z) = 0$ where $P(z) = cz^2 + (d-a)z - b$. Then multiplication by $cz + d \in \mathbb{K} = \mathbb{Q}(z)$, expressed with respect to the basis $\{z, 1\}$, is given by the transpose of the matrix g. Thus the eigenvalues of g are the $cz_j + d$ where the z_j are the roots of P. In particular 1 is not an eigenvalue of g, so that $\det(g-1) \neq 0$.
In the case $c = 0$, the equation $g(z) = z$ implies $a = d$ and $b = 0$, which gives that g sits in the diagonal subgroup $\mathbb{Q}^* \subset \mathrm{GL}_2^+(\mathbb{Q})$. Since $g \neq 1$, one has $\det(g-1) = (a-1)^2 \neq 0$. $\square$

This phenomenon of lack of injectivity of the map $\Gamma\backslash G_y \to c(p(y))$ is similar to the phenomenon of presence of holonomy in the context of foliations. As in that context, we use the Hilbert space $\ell^2(\Gamma\backslash G_y)$ instead of $\ell^2(c(p(y)))$ to construct the representation π_y.

5.3. Time evolution and regular representation.

As in the 1-dimensional case, the algebra $\mathcal{A}_2$ of commensurability classes of 2-dimensional $\mathbb{Q}$-lattices up to scale has a natural time evolution, induced by the ratio of covolumes of pairs of commensurable $\mathbb{Q}$-lattices. In terms of the description of the algebra $\mathcal{A}_2$ as a convolution algebra of functions on $\mathcal{U}_2$ as in (3.196), the time evolution is given by the one-parameter group of automorphisms σ_t,

$$\sigma_t(f)(g,\rho,z) = \det(g)^{it} f(g,\rho,z). \tag{3.205}$$

Equivalently, if $((\Lambda,\phi),(\Lambda',\phi'))$ is the pair of commensurable $\mathbb{Q}$-lattices corresponding to the data (g,ρ,z) under the map (3.187), one has

$$\det(g) = \mathrm{Covol}(\Lambda')/\mathrm{Covol}(\Lambda). \tag{3.206}$$

which is the extension of (3.58) (using (3.57)) to the two-dimensional case.
We now construct the regular representation of the algebra $\mathcal{A}_2$ as the action of the algebra $\mathcal{A}_2$ by left convolution operators on the Hilbert space $L^2(\mathcal{U}_2, dv)$ for the measure dv defined as follows. By construction $\mathcal{U}_2$ is the quotient $(\Gamma\times\Gamma)\backslash\tilde{\mathcal{U}}_2$ (cf. (3.195)). Since the action of $\Gamma\times\Gamma$ is proper, an invariant measure du on $\tilde{\mathcal{U}}_2$ uniquely determines a measure dv on the quotient whose pull-back gives du.
We let du be the measure on $\tilde{\mathcal{U}}_2 = \{(g,\rho,z)\in \mathrm{GL}_2^+(\mathbb{Q})\times M_2(\hat{\mathbb{Z}})\times\mathbb{H} \,:\, g\rho\in M_2(\hat{\mathbb{Z}})\}$ which is the product of the counting measure on $\mathrm{GL}_2^+(\mathbb{Q})$ by the measure

$$d\mu(y) = d\rho\times d\mu_{\mathbb{H}}(z), \tag{3.207}$$

on $M_2(\hat{\mathbb{Z}})\times\mathbb{H}$, where

$$d\rho = \prod d\rho_{ij}, \tag{3.208}$$

is the normalized Haar measure on the additive compact group $M_2(\hat{\mathbb{Z}})$, i.e. the product of the additive Haar measures on the column vectors, and $d\mu_{\mathbb{H}}(z)$ is given by the Riemannian volume form on $\mathbb{H}$ for the Poincaré metric, normalized so that $\mu(\Gamma\backslash\mathbb{H}) = 1$.
We denote by M the von Neumann algebra generated by $\mathcal{A}_2$ in the regular representation given by the action of $\mathcal{A}_2$ by left convolution operators on the Hilbert space $L^2(\mathcal{U}_2, dv)$.

Proposition 3.46. *The von Neumann algebra M admits the following cyclic and separating vector ξ,*

$$\xi(g,\rho,z) = 0, \quad \text{for } g\notin\Gamma, \qquad \text{and } \xi(g,\rho,z) = 1\,, \quad \forall g\in\Gamma\,. \tag{3.209}$$

Let σ_t be the time evolution of (3.205). *Then the one-parameter group of automorphisms σ_{2t} is the restriction to the C^*-algebra $\mathcal{A}_2$ of the modular automorphism group of M relative to ξ.*
The linear functional

$$\varphi(f) = \int_{\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})} f(1,\rho,z)\,d\mu(\rho,z), \tag{3.210}$$

on the C^*-algebra $\mathcal{A}_2$ is a KMS_2 state for the time evolution σ_t of (3.205).

Proof. As a measure space, the space $\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})$ with the quotient of the measure $d\mu$ of (3.207) is the total space of a bundle with base $\Gamma\backslash\mathbb{H}$ and fiber the measure space $M_2(\hat{\mathbb{Z}})/\{\pm 1\}$ with the probability measure $d\rho$, hence the total mass is

$$\int_{\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})} d\mu(\rho,z) = 1.$$

Thus the function ξ is square integrable. This function is a unit for the convolution product (3.196) and this shows that ξ is a cyclic vector for the actions of $\mathcal{A}_c(\mathcal{U}_2)$ by left and by right convolution. This suffices to show that ξ is cyclic and separating for M since the action of $\mathcal{A}_c(\mathcal{U}_2)$ by right convolution is contained in the commutant of M.

The state (3.210) satisfies

$$\begin{aligned}&\varphi(f^* * f) = \\ &\textstyle\int_{\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})} \sum_{h\in\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})\,:\,h\rho\in M_2(\hat{\mathbb{Z}})} \overline{f(h,\rho,z)} f(h,\rho,z) d\mu(\rho,z),\end{aligned} \tag{3.211}$$

for all $f\in\mathcal{A}_2$. This shows that the Hilbert space $\mathcal{H}(\varphi)$ of the GNS representation is the Hilbert space of the regular representation.

It is thus enough to prove that the state (3.210) satisfies the KMS condition at inverse temperature $\beta = 2$. Since the measure $d\rho$ is the product of the additive Haar measures on the columns as in (3.208), it transforms as

$$d(g\rho) = (\det g)^{-2} d\rho, \quad \forall g\in \mathrm{GL}_2^+(\mathbb{Q}). \tag{3.212}$$

In fact, (3.212) implies that one has

$$\begin{aligned}&\textstyle\int_{\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})} \sum_{h\in\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})} \alpha(h,\rho,z) d(\rho,z) = \\ &\textstyle\int_{\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})} \sum_{k\in\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})} \alpha(k^{-1},k\rho,k(z)) (\det k)^{-2} d(\rho,z),\end{aligned} \tag{3.213}$$

for all compactly supported continuous functions α on

$$\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q}) \times_\Gamma (M_2(\hat{\mathbb{Z}})\times\mathbb{H}).$$

We now consider, for two given elements $f_1, f_2\in\mathcal{A}_2$, a corresponding function α as above, with the property that

$$\alpha(h,\rho,z) = f_1(h^{-1},h\rho,h(z)) f_2(h,\rho,z) (\det h)^{it-2} \tag{3.214}$$

if the condition $h\rho\in M_2(\hat{\mathbb{Z}})$ is satisfied, and $\alpha(h,\rho,z) = 0$ otherwise. The left hand side of (3.213) is then equal to $\varphi(f_1\sigma_{t+2i}(f_2))$, while the right hand side gives $\varphi(\sigma_t(f_2)f_1)$. Thus, (3.214) gives the desired KMS$_2$ condition in the form

$$\varphi(f_1\sigma_{t+2i}(f_2)) = \varphi(\sigma_t(f_2)f_1). \tag{3.215}$$

This completes the proof. $\square$

5.4. Symmetries.

We now describe the natural symmetry group of the quantum statistical mechanical system $(\mathcal{A}_2,\sigma_t)$ described above.

Consider the adelic group

$$\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) = \prod_{\mathrm{res}} \mathrm{GL}_2(\mathbb{Q}_p), \tag{3.216}$$

where in the restricted product the p-component is in $\mathrm{GL}_2(\mathbb{Z}_p)$ for all but finitely many places p. Any element of the group $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ can be written as a product $g\rho$ where $g \in \mathrm{GL}_2^+(\mathbb{Q})$ and $\rho \in \mathrm{GL}_2(\hat{\mathbb{Z}})$. Thus, as subsets of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$, one has

$$\mathrm{GL}_2^+(\mathbb{Q})\mathrm{GL}_2(\hat{\mathbb{Z}}) = \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}). \tag{3.217}$$

This decomposition is not unique and the intersection of the two subgroups is $\mathrm{GL}_2^+(\mathbb{Q}) \cap \mathrm{GL}_2(\hat{\mathbb{Z}}) = \mathrm{SL}(2,\mathbb{Z})$.

The group $\mathrm{GL}_2(\hat{\mathbb{Z}})$ admits a right action on 2-dimensional $\mathbb{Q}$-lattices given by

$$(\Lambda,\phi)^\gamma = (\Lambda, \phi\circ\gamma), \quad \forall \gamma \in \mathrm{GL}_2(\hat{\mathbb{Z}}), \tag{3.218}$$

i.e. precomposing the map $\phi : \mathbb{Q}^2/\mathbb{Z}^2 \to \mathbb{Q}\Lambda/\Lambda$ by $\gamma \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ viewed as an invertible morphism $\gamma : \mathbb{Q}^2/\mathbb{Z}^2 \to \mathbb{Q}^2/\mathbb{Z}^2$. By construction, this action preserves the commensurability relation, i.e.

$$(\Lambda,\phi) \sim (\Lambda',\phi') \Leftrightarrow (\Lambda,\phi\circ\gamma) \sim (\Lambda',\phi'\circ\gamma).$$

LEMMA 3.47. *The action* (3.218) *induces the following action of* $\mathrm{GL}_2(\hat{\mathbb{Z}})$ *by automorphisms on the algebra* $\mathcal{A}_2$*:*

$$\vartheta_\gamma(f)(g,\rho,z) = f(g,\rho\gamma,z), \quad \forall f \in \mathcal{A}_2, \forall \gamma \in \mathrm{GL}_2(\hat{\mathbb{Z}}). \tag{3.219}$$

The action (3.219) *is compatible with the time evolution.*

PROOF. The action (3.218) can be written equivalently as $(\rho,z) \mapsto (\rho\gamma,z)$, for $(\rho,z) \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$. This action commutes with the partially defined left action of $\mathrm{GL}_2^+(\mathbb{Q})$, so it defines automorphisms of the algebra $\mathcal{A}_2$ as in (3.219). These clearly satisfy

$$\vartheta_{\gamma_1}\vartheta_{\gamma_2} = \vartheta_{\gamma_1\gamma_2}, \quad \forall \gamma_1,\gamma_2 \in \mathrm{GL}_2(\hat{\mathbb{Z}}). \tag{3.220}$$

The compatibility with the time evolution is evident in (3.218), since the action on a pair of commensurable $\mathbb{Q}$-lattices preserves the value of the ratio of their covolumes. $\square$

We use the notation

$$[q] := \begin{pmatrix} q & 0 \\ 0 & q \end{pmatrix}, \quad \forall q \in \mathbb{Q}^* \tag{3.221}$$

REMARK 3.48. *Note that the element* $[-1]$ *acts trivially,* $\vartheta_{[-1]} = 1$*, since it belongs to the center of* Γ*.*

The action by automorphisms of $\mathrm{GL}_2(\hat{\mathbb{Z}})$ parallels the analogous action by automorphisms of the BC system by the group $\mathrm{GL}_1(\hat{\mathbb{Z}})$. However, unlike the 1-dimensional case, in the system $(\mathcal{A}_2,\sigma_t)$ of commensurability classes of 2-dimensional $\mathbb{Q}$-lattices we encounter also important symmetries that are not given by automorphisms. These account for the $\mathrm{GL}_2^+(\mathbb{Q})$ part of the group $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$. Instead of describing separately the action of $\mathrm{GL}_2^+(\mathbb{Q})$ and then checking the compatibility with the decomposition (3.217) we shall define directly an action by endomorphisms of the semigroup

$$\Sigma = M_2(\hat{\mathbb{Z}}) \cap \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \tag{3.222}$$

Its relation with the group of symmetries $\mathbb{Q}_+^*\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ which will be shown to act on KMS states is given by:

PROPOSITION 3.49. *Let G be a group and h a semigroup morphism from Σ to G such that $h([n]) = 1$ for all $n \in \mathbb{N}$, which corresponds to the diagonal embedding $\mathbb{N} \subset M_2(\hat{\mathbb{Z}}) \cap \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$.*
Then there exists a unique extension of h to a homomorphism $\tilde{h}$ from

$$\mathbb{Q}_+^* \backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \to G\,.$$

PROOF. Let us consider the natural homomorphism

$$\iota\;:\;\Sigma = M_2(\hat{\mathbb{Z}}) \cap \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \to \mathbb{Q}_+^* \backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})\,.$$

Using (3.217) one gets that ι is surjective, since any element of $\mathrm{GL}_2^+(\mathbb{Q})$ belongs to $[n]^{-1} M_2^+(\mathbb{Z})$ for some $n \in \mathbb{N}$ while $\mathrm{GL}_2(\hat{\mathbb{Z}}) \subset \Sigma$. By construction one has $\iota([n]) = 1$ for all $n \in \mathbb{N}$. Moreover for $a, b \in \Sigma$ one has

$$\iota(a) = \iota(b) \Leftrightarrow \exists\, n_j \in \mathbb{N} \text{ such that } [n_1]\, a = [n_2]\, b\,. \tag{3.223}$$

Indeed, $\iota(a) = \iota(b)$ means that there exist two elements $q_j \in \mathbb{Q}_+^*$ with $[q_1]\, a = [q_2]\, b$ and chasing the denominators gives the required result. Thus $\mathbb{Q}_+^* \backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ is the quotient semigroup of Σ by the equivalence relation defined by (3.223). This gives the existence of the extension $\tilde{h}$ and its uniqueness comes from the surjectivity of ι. ☐

Let us now construct the action of Σ by endomorphisms of $\mathcal{A}_2$. Notice first that we can express the algebra $\mathcal{A}_2$ as the reduction by a projection of the convolution algebra of $\Gamma \times \Gamma$-invariant functions on

$$\tilde{\mathcal{W}}_2 = \mathrm{GL}_2^+(\mathbb{Q}) \times M_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}.$$

We achieve this by extending elements $f \in \mathcal{A}_2$ by zero on those (g, ρ, z) where either $\rho \notin M_2(\hat{\mathbb{Z}})$ or $g\rho \notin M_2(\hat{\mathbb{Z}})$. Then we can rewrite the convolution product (3.196) in the simpler form

$$f_1 * f_2(g, \rho, z) = \sum_{s \in \Gamma \backslash \mathrm{GL}_2^+(\mathbb{Q})} f_1(gs^{-1}, s\rho, s(z)) f_2(s, \rho, z). \tag{3.224}$$

This formula defines the convolution product in the algebra $\mathcal{B} = \mathcal{A}_c(\mathcal{W}_2)$, where $\mathcal{W}_2$ is the locally compact space obtained as the quotient

$$\mathcal{W}_2 = (\Gamma \times \Gamma) \backslash \tilde{\mathcal{W}}_2\,. \tag{3.225}$$

Let e be the idempotent multiplier given by the $(\Gamma \times \Gamma)$-invariant function:

$$e(g, \rho, z) = \begin{cases} 1 & \text{if } g \in \Gamma \text{ and } \rho \in M_2(\hat{\mathbb{Z}}) \\ 0 & \text{otherwise.} \end{cases} \tag{3.226}$$

The algebra $\mathcal{A}_c(\mathcal{U}_2)$ is the reduction $e\mathcal{B}e$ of $\mathcal{B} = \mathcal{A}_c(\mathcal{W}_2)$ by the projection e. Any automorphism $\theta \in \mathrm{Aut}(\mathcal{B})$ such that $\theta(e) \leq e$ defines by restriction an endomorphism of $\mathcal{A}_c(\mathcal{U}_2)$. Moreover the map

$$\theta \in \mathrm{Aut}(\mathcal{B})\,,\ \theta(e) \leq e \mapsto \theta|_{e\mathcal{B}e} \tag{3.227}$$

is a homomorphism from the semigroup $\{\theta \in \mathrm{Aut}(\mathcal{B})\,,\ \theta(e) \leq e\}$ to the endomorphisms of $e\mathcal{B}e$. Exactly as in Lemma 3.47 one gets

LEMMA 3.50. *The following defines an action of* $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ *by automorphisms of* $\mathcal{B}$,

$$\theta_\gamma(f)(g,\rho,z) = f(g,\rho\gamma,z), \quad \forall f \in \mathcal{B}, \forall \gamma \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}). \tag{3.228}$$

Let $\gamma \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$. *If* $\gamma^{-1} \in M_2(\hat{\mathbb{Z}})$ *one has* $\theta_\gamma(e) \le e$.

PROOF. The proof of the first statement is as in Lemma 3.47. One has

$$\theta_\gamma(e)(g,\rho,z) = \begin{cases} 1 & \text{if } g \in \Gamma \text{ and } \rho\gamma \in M_2(\hat{\mathbb{Z}}) \\ 0 & \text{otherwise.} \end{cases}$$

but $\rho\gamma \in M_2(\hat{\mathbb{Z}})$ implies that $\rho \in M_2(\hat{\mathbb{Z}})$ since $\gamma^{-1} \in M_2(\hat{\mathbb{Z}})$. Thus, one obtains $\theta_\gamma(e) \le e$. □

Consider now elements $m \in M_2(\hat{\mathbb{Z}}) \cap \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$. For $m \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ we have

$$\det(m) \in \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},f}) = \mathbb{Q}_+^* \times \mathrm{GL}_1(\hat{\mathbb{Z}}), \tag{3.229}$$

where the right hand side is the product of the corresponding groups. We denote by $\det_{\mathbb{Q}}(m)$ the $\mathbb{Q}_+^*$ component of $\det(m)$. For $m \in M_2(\hat{\mathbb{Z}}) \cap \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ this component satisfies $\det_{\mathbb{Q}}(m) \in \mathbb{N} = \mathbb{Z}_{>0}$, since it belongs to $\mathbb{Q}_+^* \cap \hat{\mathbb{Z}}$.

LEMMA 3.51. *The map* $\nu(m) := m \det_{\mathbb{Q}}(m)^{-1}$ *is a semigroup homomorphism from* Σ *to*

$$\{\gamma \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \,,\ \gamma^{-1} \in M_2(\hat{\mathbb{Z}})\}.$$

PROOF. The map $m \mapsto \det_{\mathbb{Q}}(m)$ is a homomorphism from Σ to the center of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ (we omit for simplicity the brackets, i.e. the map $n \mapsto [n]$). This shows that ν is a semigroup homomorphism. It remains to show that for $m \in \Sigma$ one has $\nu(m)^{-1} \in M_2(\hat{\mathbb{Z}})$, i.e. that $(m \det_{\mathbb{Q}}(m)^{-1})^{-1}$ belongs to $M_2(\hat{\mathbb{Z}})$. In fact, we have

$$\tilde{m} := m^{-1}\det_{\mathbb{Q}}(m) \in M_2(\hat{\mathbb{Z}}), \quad \forall m \in \Sigma. \tag{3.230}$$

Indeed, we have $\tilde{m} = m^{-1}\det_{\mathbb{Q}}(m) = m^{-1}\det(m)\alpha$, for some $\alpha \in \mathrm{GL}_1(\hat{\mathbb{Z}})$, while $m^{-1}\det(m) \in M_2(\hat{\mathbb{Z}})$, since it is given by the matrix

$$m^{-1}\det(m) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{for} \quad m = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

□

PROPOSITION 3.52. *The equality*

$$\vartheta_m = \theta_{\nu(m)}|_{\mathcal{A}_2} \tag{3.231}$$

defines an action of Σ *by endomorphisms of the algebra* $\mathcal{A}_2$ *such that*

$$\vartheta_m(f)(g,\rho,z) = \begin{cases} f(g,\rho m \det_{\mathbb{Q}}(m)^{-1},z) & \text{if } \rho m \det_{\mathbb{Q}}(m)^{-1} \in M_2(\hat{\mathbb{Z}}) \\ & \text{and } g\rho m \det_{\mathbb{Q}}(m)^{-1} \in M_2(\hat{\mathbb{Z}}) \\ 0 & \text{otherwise.} \end{cases} \tag{3.232}$$

PROOF. Combining Lemma 3.51 and 3.50 with (3.227) gives that (3.231) defines an action of Σ by endomorphisms of the reduced algebra $e\mathcal{B}e$ and hence of $\mathcal{A}_2$. The equality 3.232 follows from (3.228) since elements $f \in \mathcal{A}_2$ are extended by zero on those (g,ρ,z) where either $\rho \notin M_2(\hat{\mathbb{Z}})$ or $g\rho \notin M_2(\hat{\mathbb{Z}})$. □

By construction ϑ_m is an isomorphism with the reduced algebra by the projection $\theta_{\nu(m)}(e)$. We let R_m be the range of right multiplication by $\tilde{m}$ on $M_2(\hat{\mathbb{Z}})$. The condition $\rho\tilde{m}^{-1} \in M_2(\hat{\mathbb{Z}})$ is equivalent to $\rho \in M_2(\hat{\mathbb{Z}})\tilde{m} = R_m$. By construction R_m is an open and closed subset of $M_2(\hat{\mathbb{Z}})$ which is invariant under left multiplication by Γ. We can thus write explicitly $\theta_{\nu(m)}(e)$ as the function

$$\theta_{\nu(m)}(e)(g,\rho,z) = \begin{cases} 0 & g \notin \Gamma \\ \mathbf{1}_{R_m}(\rho) & g \in \Gamma, \end{cases} \tag{3.233}$$

where $\mathbf{1}_{R_m}$ is the characteristic function of $R_m \subset M_2(\hat{\mathbb{Z}})$. We let

$$e_{m(\mathbb{Z}^2)}(\rho,z) = \mathbf{1}_{R_m}(\rho)\,, \quad \forall (\rho,z) \in \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})\,. \tag{3.234}$$

It is the characteristic function of

$$\Gamma\backslash(R_m \times \mathbb{H}) \subset \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H}).$$

Equivalently, this is the characteristic function of the subset of $\mathbb{Q}$-lattices of the form $(\Lambda, \phi \circ \tilde{m})$.

To understand the $\theta_{\nu(m)}(e)$ and the $e_{m(\mathbb{Z}^2)}$, we can use the decomposition (3.217) to reduce to the case of $m \in M_2^+(\mathbb{Z})$, where $M_2^+(\mathbb{Z})$ is the semigroup of integral matrices with positive determinant. In fact, for $m \in \Sigma$, we can write $m = m_0\alpha_0$ for $m_0 \in M_2^+(\mathbb{Z})$ and $\alpha_0 \in \mathrm{GL}_2(\hat{\mathbb{Z}})$. One has $\theta_{\nu(m)}(e) = \theta_{\nu(m_0)}(e)$ and $e_{m(\mathbb{Z}^2)} = e_{m_0(\mathbb{Z}^2)}$. For $m \in M_2^+(\mathbb{Z})$ we have

$$\tilde{m} = \det(m)\, m^{-1} \in M_2^+(\mathbb{Z}). \tag{3.235}$$

The notation $e_{m(\mathbb{Z}^2)}$ suggests that it only depends upon the range $m(\mathbb{Z}^2)$. This is justified by the following result.

LEMMA 3.53. *Let $m \in M_2^+(\mathbb{Z})$. The range $R_m \subset M_2(\hat{\mathbb{Z}})$ of right multiplication by $\tilde{m}$ and the projection $e_{m(\mathbb{Z}^2)}$ only depend on the subgroup $m(\mathbb{Z}^2) \subset \mathbb{Z}^2$.*

PROOF. If two elements $m_1, m_2 \in M_2^+(\mathbb{Z})$ satisfy $m_1(\mathbb{Z}^2) = m_2(\mathbb{Z}^2)$, then $m_2 = m_1\gamma$ for some $\gamma \in \Gamma = \mathrm{SL}_2(\mathbb{Z})$. Thus, one obtains

$$M_2(\hat{\mathbb{Z}})\tilde{m}_1 = M_2(\hat{\mathbb{Z}})\tilde{m}_2.$$

$\square$

These projections satisfy the relations

$$e_{m(\mathbb{Z}^2)}e_{m'(\mathbb{Z}^2)} = e_{m(\mathbb{Z}^2)\cap m'(\mathbb{Z}^2)}, \qquad e_{\mathbb{Z}^2} = 1. \tag{3.236}$$

Let $m \in M_2^+(\mathbb{Z})$ and $L = (\Lambda, \phi)$ be a $\mathbb{Q}$-lattice in the support of the characteristic function $e_{m(\mathbb{Z}^2)}$. Then we can define

$$L \circ \tilde{m}^{-1} := (\Lambda, \phi \circ \tilde{m}^{-1}). \tag{3.237}$$

This transformation preserves commensurability and it corresponds to the endomorphism ϑ_m using (3.235).

In order to describe how the symmetries will act on KMS states of the GL_2-system $(\mathcal{A}_2, \sigma_t)$, we need to identify which part of the symmetries described above is *inner*.

Consider for $n \in \mathbb{N}$ the isometry

$$\mu_{[n]}(g,\rho,z) = \begin{cases} 1 & g \in \Gamma \cdot [n] \\ 0 & \text{otherwise.} \end{cases} \tag{3.238}$$

The elements $\mu_{[n]}$ are multipliers of the algebra $\mathcal{A}_2$,

$$\mu_{[n]} \in M(\mathcal{A}_2). \tag{3.239}$$

LEMMA 3.54. *The $\mu_{[n]}$ satisfy*

$$\mu_{[n]}^* * \mu_{[n]} = 1, \tag{3.240}$$

while

$$\mu_{[n]} * \mu_{[n]}^* = \pi_n, \tag{3.241}$$

where π_n is the characteristic function of the set of $\mathbb{Q}$-lattices divisible by n. Moreover, the $\mu_{[n]}$ are eigenvectors of the time evolution, namely they satisfy

$$\sigma_t(\mu_{[n]}) = n^{2it}\mu_{[n]}. \tag{3.242}$$

PROOF. One has

$$\mu_{[n]}^*(g,\rho,z) = \begin{cases} 1 & g \in \Gamma \cdot [n]^{-1} \ \text{ and } \ n \mid \rho \\ 0 & \text{otherwise.} \end{cases} \tag{3.243}$$

The result then follows by a direct calculation. □

We have then the following result about the endomorphisms $\vartheta_{[n]}$.

PROPOSITION 3.55. *For all $n \in \mathbb{N} \subset M_2^+(\mathbb{Z})$, the endomorphism $\vartheta_{[n]}$ of $\mathcal{A}_2$ is implemented by the multiplier $\mu_{[n]}$ in the form*

$$\vartheta_{[n]}(f) = \mu_{[n]}\, f\, \mu_{[n]}^*, \qquad \forall f \in \mathcal{A}_2. \tag{3.244}$$

PROOF. Both endomorphisms $\theta_{[n]}$ and $\mu_{[n]}(\cdot)\mu_{[n]}^*$ have range the reduced algebra by $e_{[n]}$. Thus one needs to evaluate $\mu_{[n]} * f * \mu_{[n]}^*(g,\rho,z)$ assuming $n \mid g\rho$ and $n \mid \rho$. It is given by

$$\sum_{h \in \Gamma \backslash \mathrm{GL}_2^+(\mathbb{Q})} \mu_{[n]}(gh^{-1}, h\rho, h(z))(f * \mu_{[n]}^*)(h,\rho,z), \tag{3.245}$$

where the expression can be nonzero only for the value

$$h = [n]^{-1}g \in \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q}).$$

For this value we have $h\rho \in M_2(\hat{\mathbb{Z}})$ since we are assuming that $g\rho$ is divisible by n. Thus, we get

$$\begin{aligned} \mu_{[n]} * f * \mu_{[n]}^*(g,\rho,z) = & \ (f * \mu_{[n]}^*)([n]^{-1}g,\rho,z) \\ = & \sum_{s \in \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})} f([n^{-1}]gs^{-1}, s\rho, s(z))\mu_{[n]}^*(s,\rho,z). \end{aligned}$$

Here the only possible value for s is $[n]^{-1} \in \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$, so that we obtain

$$\mu_{[n]} * f * \mu_{[n]}^*(g,\rho,z) = f(g, \rho[n]^{-1}, z),$$

which is the same as $\vartheta_{[n]}(f)(g,\rho,z)$. □

Combining the results of this section we can identify the following symmetry group that acts on KMS states of the system $(\mathcal{A}_2, \sigma_t)$.

THEOREM 3.56. *Proposition 3.52 gives an action of the group*

$$\mathbb{Q}^* \backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \tag{3.246}$$

as symmetries of KMS states of the system $(\mathcal{A}_2, \sigma_t)$.

PROOF. We have seen in Section 2.3 above that symmetries of KMS states are given by the symmetries (by automorphisms or endomorphisms) of the system $(\mathcal{A}_2, \sigma_t)$ modulo the inner ones. Thus, Proposition 3.49, using for G the group of transformations of KMS states, shows that one gets an action of the group $\mathbb{Q}^*_+ \backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$. But Remark 3.48 shows that it passes to the quotient by $-1 \in \mathbb{Q}^*$. $\square$

We are going to describe in Section 7.4 below the explicit action of the symmetries (3.246) on KMS states of $(\mathcal{A}_2, \sigma_t)$. In particular, one sees then that the action is nontrivial on (some) KMS states, so that indeed the only inner symmetries are the ones described here in Proposition 3.55.

6. The modular field

Before discussing the arithmetic properties of the GL_2-system $(\mathcal{A}_2, \sigma_t)$, we need to review some background material on the field of modular functions (cf. e.g. [**268**], [**199**]). For each positive integer N the congruence subgroup $\Gamma(N)$ is defined as

$$\Gamma(N) = \{\gamma \in \Gamma = \mathrm{SL}_2(\mathbb{Z}) \,:\, \gamma \equiv 1 \mod N\}\,. \tag{3.247}$$

We use the standard 'slash operator' notation for the action of $\mathrm{GL}_2^+(\mathbb{R})$ on functions on the upper half plane,

$$f|_k\alpha(z) = \det(\alpha)^{k/2} f\left(\frac{az+b}{cz+d}\right)(cz+d)^{-k}, \tag{3.248}$$

with

$$\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R})\,.$$

6.1. The modular field of level $N = 1$.

We begin by describing modular functions of level one. Let $\Gamma\backslash\mathbb{H}$ be the quotient of the upper half plane $\mathbb{H}$ by the action of $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. We write the quotient map as $p : \mathbb{H} \to \Gamma\backslash\mathbb{H}$. Modular functions of level one are functions on $\Gamma\backslash\mathbb{H}$ such that

$$f \circ p \quad \text{is a holomorphic function on } \mathbb{H} \tag{3.249}$$

and such that the estimate

$$|f(p(z))| \le C e^{k\Im z} \tag{3.250}$$

holds for some $C < \infty$ and $k < \infty$ and for $\Im z \to +\infty$.
Let then $\tilde{f}(q) = f(p(z))$ where $q = e^{2\pi i z}$. This is well defined for $|q| < 1$, $q \neq 0$ and gives a holomorphic function of q in the punctured disk $D^* = D \smallsetminus \{0\}$. Condition

(3.250) shows that $q^n \tilde{f}(q)$ is bounded near $q = 0$ for some n and thus that $\tilde{f}(q)$ has an expansion of the form

$$\tilde{f}(q) = a_{-n}q^{-n} + a_{-(n-1)}q^{-(n-1)} + \cdots + a_0 + \sum_{n=1}^{\infty} a_n q^n. \tag{3.251}$$

Notice that if a modular function f is bounded then it is constant. In fact, the quotient $\Gamma(2)\backslash\mathbb{H}$ is the Riemann sphere $\mathbb{P}^1(\mathbb{C})$ with three points removed and any bounded function holomorphic on $\mathbb{P}^1(\mathbb{C})\backslash\{$ three points $\}$ is constant. This implies that the pullback of f to $\mathbb{H}$ is constant, hence so is f.

The fundamental example of a modular function of level one is the Jacobi modular function j. In fact, over $\mathbb{C}$ the ring of modular functions of level one can be identified with the ring of polynomials $\mathbb{C}[j]$. The field of fractions is called the modular field of level 1 (over $\mathbb{C}$) and is identified with the field of rational functions $\mathbb{C}(j)$.
The function $j(\tau)$ is obtained as follows (cf. [**262**]). Given a lattice $\Lambda \subset \mathbb{C}$ and $z \notin \Lambda$ one lets, with $\Lambda^* = \Lambda \smallsetminus \{0\}$,

$$\wp_\Lambda(z) = \frac{1}{z^2} + \sum_{\gamma\in\Lambda^*} \left(\frac{1}{(z+\gamma)^2} - \frac{1}{\gamma^2} \right). \tag{3.252}$$

One has

$$(z+\gamma)^{-2} - \gamma^{-2} = \sum_{n=1}^{\infty} (-1)^n (n+1) \frac{z^n}{\gamma^{n+2}} \tag{3.253}$$

so that the Taylor expansion of $\wp(z)$ near $z = 0$ is of the form

$$\wp_\Lambda(z) = \frac{1}{z^2} + \sum_{k=1}^{\infty} (2k+1) e_{2k+2}(\Lambda) z^{2k}, \tag{3.254}$$

where the coefficients $e_{2k}(\Lambda)$ are given by

$$e_{2k}(\Lambda) = \sum_{\gamma\in\Lambda^*} \gamma^{-2k}. \tag{3.255}$$

Comparing the pole part of the two sides then yields the equality

$$\wp'^2 = 4\wp^3 - g_2\wp - g_3, \quad \text{with} \quad g_2 = 60e_4, \quad \text{and} \quad g_3 = 140e_6. \tag{3.256}$$

The discriminant of the equation $4x^3 - g_2x - g_3 = 0$ is given by $(\alpha-\beta)^2(\beta-\gamma)^2(\gamma-\alpha)^2$, where α, β, and γ are the roots, and one gets for its value

$$\frac{1}{16}(g_2^3 - 27g_3^2), \tag{3.257}$$

using the symmetric functions

$$\alpha + \beta + \gamma = 0, \quad \alpha\beta + \beta\gamma + \gamma\alpha = \frac{-g_2}{4}, \quad \text{and} \quad \alpha\beta\gamma = \frac{g_3}{4}.$$

It is more convenient to use the normalization,

$$\Delta = g_2^3 - 27g_3^2. \tag{3.258}$$

One shows that $\mathbb{C}/\Lambda \ni z \mapsto (\wp(z), \wp'(z))$ is an isomorphism with the elliptic curve given by the projective algebraic curve of equation $y^2 = 4x^3 - g_2 x - g_3$. Thus, the discriminant Δ does not vanish. One then lets

$$j = 1728 \frac{g_2^3}{\Delta}. \tag{3.259}$$

We now recall how to compute its q-expansion. For $m > 1$, $q = e^{2\pi i z}$, one has

$$\sum_{n\in\mathbb{Z}} (n+z)^{-m} = (-1)^m \frac{(2i\pi)^m}{(m-1)!} \left(q \frac{\partial}{\partial q}\right)^{m-1} \left(\frac{1}{1-q}\right), \tag{3.260}$$

which is obtained by differentiating $\sum_{n\in\mathbb{Z}}(n+z)^{-1} = \pi \cot \pi z = i\pi \frac{q+1}{q-1}$. We then view e_{2k} as a function of τ, by applying it to the lattice $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$ and dividing then by $(2\pi i)^{2k}$ (which is the same as using the lattice $2\pi i(\mathbb{Z} + \mathbb{Z}\tau)$. The double sum (3.255) can then be expressed as

$$e_{2k}(\tau) = (2\pi i)^{-2k}\, 2\,\zeta(2k) + 2 \sum_{n=1}^{\infty} R_{2k}(q^n) \tag{3.261}$$

with

$$R_m(q) = \frac{(-1)^m}{(m-1)!} \left(q \frac{\partial}{\partial q}\right)^{m-1} \left(\frac{1}{1-q}\right). \tag{3.262}$$

One has

$$(m-1)!\, R_m(q) = (-1)^m \sum_{n=1}^{\infty} n^{m-1} q^n, \tag{3.263}$$

so that the q-expansion of e_{2k} is given by (cf. [**262**] VII, Proposition 8):

$$e_{2k} = (-1)^k \frac{B_k}{(2k)!} + \frac{2}{(2k-1)!} \sum_{m=1}^{\infty} \sigma_{2k-1}(m) q^m, \tag{3.264}$$

where $\sigma_r(m) = \sum_{d|m} d^r$ and the B_k are the Bernoulli numbers. With this normalization one obtains

$$\Delta = q - 24\,q^2 + 252\,q^3 - 1472\,q^4 + 4830\,q^5 + \cdots \tag{3.265}$$

and

$$j = \frac{1}{q} + 744 + 196884\,q + 21493760\,q^2 + \cdots \tag{3.266}$$

We can now recall the following result about modular functions of level one.

PROPOSITION 3.57. *Let f be a modular function of level one as above. Then the following conditions are equivalent.*

a) *All the coefficients a_n of (3.251) satisfy $a_n \in \mathbb{Q}$.*
b) *All a_n for $n \leqq 0$ belong to $\mathbb{Q}$.*
c) *There exists a polynomial P with rational coefficients such that $f = P(j)$, where j is the Jacobi modular function.*

PROOF. We have seen that $j(z)$ is a modular function with

$$\tilde{j}(q) = \frac{1}{q} + \sum_{n=0}^{\infty} j_n q^n, \quad \text{with } j_n \in \mathbb{Q}, \quad \forall n \in \mathbb{N} \cup \{0\}. \tag{3.267}$$

Clearly a) $\Rightarrow$ b). Assuming b) we take $f - a_{-n} j^n$. Then the order of the pole in q is at most $n-1$. Thus, we can find P by induction, noting that $f - a_{-n} j^n$ still fulfills b). We then get $f - P(j)$ having no pole part. It is then bounded and must be constant. Moreover this constant is rational. Thus b) $\Rightarrow$ c) and since c) $\Rightarrow$ a) we get the equivalence. $\square$

6.2. Modular field of level N.

We first work over $\mathbb{C}$ and define modular functions of level N as holomorphic functions of $\tau \in \mathbb{H}$ such that

$$f \mid \gamma = f, \quad \forall \gamma \in \Gamma(N) = \{\gamma \in \Gamma \mid \gamma \equiv 1 \mod N\} \tag{3.268}$$

and such that the estimate

$$|\,(f \mid \gamma)(\tau)\,| \le C\, e^{k\, \Im \tau} \tag{3.269}$$

holds for all $\gamma \in \Gamma(N) \backslash \Gamma$ and for $\Im\tau \to \infty$.
These functions form an algebra over $\mathbb{C}$ and the field of quotients $F_N(\mathbb{C})$ is obtained by replacing the condition that f is holomorphic with the requirement that it be meromorphic with a finite number of poles in $\Gamma(N)\backslash\mathbb{H}$.
One can compactify the quotient $\Gamma(N)\backslash\mathbb{H}$ by adding finitely many cusps (cf. [**268**] §1.5). The cusps are given by the points of the finite set

$$\Gamma(N)\backslash\mathbb{P}_1(\mathbb{Q}). \tag{3.270}$$

One sets $\mathbb{H}^* = \mathbb{H} \cup \mathbb{P}_1(\mathbb{Q})$ and obtains a compact Riemann surface as the quotient $\Sigma_N = \Gamma(N)\backslash\mathbb{H}^*$. By construction, $F_N(\mathbb{C})$ is the field of meromorphic functions on the Riemann surface Σ_N.
The group $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ acts on $F_N(\mathbb{C})$ on the right by

$$f \mapsto f \mid \gamma =: f^\gamma, \quad \forall f \in F_N(\mathbb{C}). \tag{3.271}$$

Thus, we have a finite group of automorphisms $H_N = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ of the field $F_N(\mathbb{C})$ and the fixed field $F_N(\mathbb{C})^{H_N}$ is just the field $\mathbb{C}(j)$ of meromorphic modular functions of level one. Thus by a general fact of Galois theory (Appendix Proposition 4.212) we obtain the following result (cf. [**268**]).

THEOREM 3.58. *The field $F_N(\mathbb{C})$ is a finite Galois extension of $\mathbb{C}(j)$ with Galois group $H_N = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$.*

We now recall the construction of explicit generators $\mathfrak{f}_u$ of $F_N(\mathbb{C})$ over $\mathbb{C}(j)$ using the coordinates of torsion points of an elliptic curve with j-invariant $j(\tau) = j$. One first rescales x and y in the equation $y^2 = 4x^3 - g_2 x - g_3$ as follows

$$X = \lambda\, x, \quad Y = \frac{1}{2}\, \lambda^{3/2}\, y, \quad \text{and} \quad \lambda = -2^7\, 3^5\, g_2\, g_3/\Delta, \tag{3.272}$$

so that the elliptic curve is defined over the field $\mathbb{Q}(j)$ by the equation

$$Y^2 = X^3 - 3\, j\, (j - 1728)\, X + 2\, j\, (j - 1728)^2 \tag{3.273}$$

and the discriminant is $186624\,(j-1728)^3\,j^2$. In fact this rescaling only works if $g_2\,g_3 \neq 0$ and becomes singular at the two points $g_2 = 0$, where $j = 0$ and $g_3 = 0$, where $j = 1728$.

We consider the elliptic curve $E(j)$ given by (3.273), with origin the triple point $(X, Y, Z) = (0, 1, 0)$ in homogeneous coordinates. The opposite of $(X, Y) = (X, Y, 1)$ is $(X, -Y)$, with respect to the group law where the addition is computed using

$$A + B + C = 0$$
$$\Updownarrow$$
$$(A, B, C) \text{ on the same line}$$

Let $\tau \in \mathbb{H}$ be such that $j(\tau) = j$ and $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$ the corresponding lattice. The map $z \in \mathbb{C}/\Lambda \mapsto (\lambda\wp(z), \frac{1}{2}\,\lambda^{3/2}\wp'(z))$ is an isomorphism $\mathbb{C}/\Lambda \to E(j)$.

Then, given N, one considers the labeling of the N-torsion points of $\mathbb{C}/\Lambda$ given by

$$\lambda_{(a,b)} = a - b\,\tau \in \frac{1}{N}\,\Lambda/\Lambda \tag{3.274}$$

for all $u = (a, b) \in (\frac{1}{N}\mathbb{Z}/\mathbb{Z})^2$. The function $\mathfrak{f}_u$ gives the X coordinate of the torsion point of $E(j)$ associated to λ_u, $u \neq 0$.

We start by writing the q-expansion of weight two Eisenstein series, which give the values of the $\wp$-function at torsion points. We use (3.260) to get, for $\Im u > 0$,

$$\sum \frac{1}{(n+u)^2} = (2\pi i)^2 \sum_{n=1}^{\infty} n\, e^{2\pi i n u}. \tag{3.275}$$

We divide the Eisenstein series by $(2\pi i)^2$, which gives, with $q = e^{2\pi i\tau}$, $z = e^{2\pi i u}$,

$$\begin{aligned}
&(2\pi i)^{-2}\left(\sum_{n,m} \frac{1}{(n+m\tau+u)^2} - \sum_{(n,m)\neq(0,0)} \frac{1}{(n+m\tau)^2}\right) \\
=\ & \frac{z}{(1-z)^2} + \sum_{m>0}\sum_{n>0} n(z^n + z^{-n})\,q^{nm} - \left((2\pi i)^{-2}\,\frac{\pi^2}{3} + 2\sum_{m>0}\sum_{n>0} n\,q^{nm}\right) \\
=\ & \frac{1}{12} + \frac{z}{(1-z)^2} + \sum_{d|n,n} d(z^d + z^{-d} - 2)\,q^n\,.
\end{aligned}$$

In general, for $(a, b) \in \{0, 1, \ldots, N-1\}^2 \neq (0, 0)$, we let $z = e^{2\pi i \frac{a}{N}}\,q^{\frac{-b}{N}} \neq 1$ and define

$$E_{(\frac{a}{N},\frac{b}{N})}(q) = \frac{1}{12} + \frac{z}{(1-z)^2} + \sum_{n,d|n} d(z^d + z^{-d} - 2)\,q^n, \tag{3.276}$$

and the Fricke functions

$$\mathfrak{f}_{(\frac{a}{N},\frac{b}{N})}(\tau) = -2^7\,3^5\,\frac{g_2\,g_3}{\Delta}\,E_{(\frac{a}{N},\frac{b}{N})}(q)\,, \quad q = e^{2\pi i\tau}\,. \tag{3.277}$$

PROPOSITION 3.59. *The function $\mathfrak{f}_u$ is in $F_N(\mathbb{C})$ and its q-expansion has all its coefficients in $\mathbb{Q}(e^{2\pi i/N})$. The $\mathfrak{f}_u$, $u \in (\frac{1}{N}\mathbb{Z}/\mathbb{Z})^2$, $u \neq 0$, generate $F_N(\mathbb{C})$ over $\mathbb{C}(j)$.*

PROOF. Since $b \in \{0, 1, \ldots, N-1\}$, the term in $z^d\,q^n$ contains a positive power of q even for $d = n$. Thus, the series (3.276) makes sense first as a formal power series with coefficients in $\mathbb{Q}(e^{2\pi i/N})$. In fact, it is a series in $q^{1/N}$ and the coefficient

of the n-th term is $O(n^t)$, so that it is convergent for $|q| < 1$. The fact that $E_{(\frac{a}{N},\frac{b}{N})}$ is a modular form of weight 2 and level N follows from the way $\mathrm{SL}_2(\mathbb{Z})$ acts through its action on the labels (a, b). This makes it possible to check the behavior at all cusps and the $\Gamma(N)$-modularity. Once multiplied as in (3.277), one gets a modular function with no pole in $\mathbb{H}$ since Δ has no zero there. Also, the q-expansion still has coefficients in $\mathbb{Q}(e^{2\pi i/N})$.
That the $\mathfrak{f}_u$ generate $F_N(\mathbb{C})$ over $\mathbb{C}(j)$ follows from Theorem 4.215 of the appendix, and the fact that no element $g \in H_N$, $g \neq 1$ acts trivially on the $\mathfrak{f}_u$ since it acts by a non-trivial permutation of the labels (a, b). □

We now define and analyze the modular field F_N of level N as a subfield of $F_N(\mathbb{C})$.

Definition 3.60. *Let $F_N \subset F_N(\mathbb{C})$ be the subfield generated over $\mathbb{Q}$ by j and the $\mathfrak{f}_u$ for $u \in (\frac{1}{N}\mathbb{Z}/\mathbb{Z})^2$, $u \neq 0$.*

Notice that, if N is even, we do not need j, since at level 2 it is a simple function of the $\mathfrak{f}_u$, since the X coordinates of 2-division points are the zeros of Y and their symmetric functions give j. In fact, one has

$$j = 256(1 - f + f^2)^3/(f^2(1 - f)^2),$$

where (cf. (3.296) below)

$$f = (\mathfrak{f}_{(\frac{1}{2},\frac{1}{2})} - \mathfrak{f}_{(0,\frac{1}{2})})/(\mathfrak{f}_{(\frac{1}{2},0)} - \mathfrak{f}_{(0,\frac{1}{2})}).$$

By construction F_N is a countable field which contains the field $\mathbb{Q}(j)$ of rational functions of j as a subfield. The Galois group $H_N = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ of $F_N(\mathbb{C})/\mathbb{C}(j)$ leaves F_N globally invariant since it fixes j and permutes the $\mathfrak{f}_u$. One might think at first sight that the Galois group of $F_N/\mathbb{Q}(j)$ will be the same as for $F_N(\mathbb{C})/\mathbb{C}(j)$ but it is finer in that $H_N = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ gets replaced by $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$. This is easy to understand in terms of the automorphism group of the abelian group of N-torsion points of elliptic curves (cf. [**268**]), as briefly recalled below. The main point is the determination of the fixed field $F_N^{H_N}$ as the field $\mathbb{Q}(j, e^{2\pi i/N})$ and one needs in particular to show that F_N contains a primitive N-th root of unity (cf. [**268**] Theorem 6.6). We do this concretely, using θ-functions, i.e. after identifying $E(j)_{\mathbb{C}}$ with $\mathbb{C}/\Lambda$ with $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$, we use θ-functions to write explicitly an elliptic function with a given divisor. We first recall the following standard properties of θ-functions.

Lemma 3.61. *Let $q = e^{2\pi i \tau}$ and set*

$$\theta_1(z) = 2\, q^{1/8} \sum_{k=0}^{\infty} (-1)^k\, q^{\frac{k(k+1)}{2}} \sin((2k+1)\pi z). \tag{3.278}$$

Then the following hold.

- $\theta_1(z + 1) = -\theta_1(z)$
- $\theta_1(z + \tau) = -q^{-1/2}\, e^{-2\pi i z}\, \theta_1(z)$
- *$\theta_1(z)$ vanishes exactly for $z \in \Lambda$ with simple zeros.*

Proof. The first two properties are easy to check. For the third property one uses the product formula

$$\theta_1(z) = 2\, \lambda \sin \pi z \prod_{k=1}^{\infty} (1 - q^k \cos 2\pi z + q^{2k}) \tag{3.279}$$

with $\lambda = q^{1/8}\prod_{k=1}^{\infty}(1-q^k)$. This makes it possible to check the property of the zeros directly. $\square$

Then, for any t with $N^2\,t\in\mathbb{Z}$, one lets, with the notation (3.274)

$$E_N^0 = \{\lambda_{(\frac{a}{N},\frac{b}{N})}\,|\,a,b\in\{0,\ldots,N-1\}\}\,,$$

$$h_t(z) = \prod_{v\in E_N^0}\frac{\theta_1(t+z+v)}{\theta_1(z+v)}. \tag{3.280}$$

LEMMA 3.62. *The function $h_t(z)$ of (3.280) satisfies the following properties.*

(1) *$h_t(z)$ is an elliptic function of $z\in E=\mathbb{C}/\Lambda$.*
(2) *Let $\alpha = Nt$. Then h_t is the unique (up to multiplication by $\mathbb{C}^*$) elliptic function on E which has simple zeros on $\{z\in E\,,\;Nz=\alpha\}$ and simple poles on $E_N=\frac{1}{N}\Lambda/\Lambda$.*
(3) *For all $s\in\frac{1}{N}\Lambda$ one has $h_t(z+s)=\omega\,h_t(z)$, where ω is an N-th root of 1.*
(4) *One has*

$$h_{1/N^2}(z-\tau/N) = e^{2\pi i/N}\,h_{1/N^2}(z)\,. \tag{3.281}$$

PROOF. (1) First notice that $h_t(z+1)=h_t(z)$. Next, for $h_t(z+\tau)$, one gets for each of the terms in the product a multiplicative factor or $e^{-2\pi it}$. Thus, one gets an overall factor of $(e^{-2\pi it})^{N^2}=1$, since $N^2\,t\in\mathbb{Z}$.
(2) By Lemma 3.61 the formula (3.280) shows that h_t satisfies the stated condition on zeros and poles. The ratio of any two elliptic functions fulfilling these conditions is an elliptic function without pole and hence a constant.
(3) By the above the action of translations by the group E_N admits h_t as an eigenfunction and the associated character $\chi: E_N\to\mathbb{C}^*$ takes its values in the group of N-th roots of 1.
(4) We fix $t=\frac{1}{N^2}$ and $s=-\frac{1}{N}\tau$. Using the notation (3.274) one has $\lambda_{(a,\frac{b}{N})}+s=\lambda_{(a,\frac{b+1}{N})}$ for $b<N-1$ and $\lambda_{(a,\frac{N-1}{N})}+s=\lambda_{(a,0)}-\tau$. Thus $h_t(z+s)/h_t(z)$ simplifies except for the boundary terms in (3.280) involving the transition $\lambda_{(a,\frac{N-1}{N})}+s=\lambda_{(a,0)}-\tau$. This gives a multiplicative factor

$$\prod_a e^{2\pi i(t+z+\lambda_{(a,0)})}/\prod_a e^{2\pi i(z+\lambda_{(a,0)})} = e^{2\pi iNt} = e^{2\pi i/N}\,.$$

which proves (3.281). $\square$

Recall now that $j:\Gamma\backslash\mathbb{H}\to\mathbb{C}$ is a bijection (cf. [**262**]). We let $\tau\in\mathbb{H}$ be such that the value $j=j(\tau)$ is a transcendental number. The field inclusion $\mathbb{Q}(j)\subset\mathbb{C}$ is well defined. For a given $N\in\mathbb{N}$, we then let $F_N(j)$ be the subfield of $\mathbb{C}$ generated by the coordinates X of the (non zero) N-torsion points $(X,Y,1)$ of the elliptic curve $E(j)$.

LEMMA 3.63. *The field $F_N(j)$ is a finite Galois extension of $\mathbb{Q}(j)$ and its Galois group injects in the quotient by ±1 of the group of automorphisms of the abelian group of N-torsion points of $E(j)$.*

PROOF. First, using the $\wp$- and $\wp'$-functions to identify $E(j)_{\mathbb{C}}$ with $\mathbb{C}/\Lambda$, we get N^2 points P of $E(j)_{\mathbb{C}}$ such that $N \cdot P = 0$ (N-torsion points). Let $\sigma \in \mathrm{Aut}_{\mathbb{Q}(j)}\mathbb{C}$ be any automorphism of $\mathbb{C}$ fixing j. The action of σ on the coordinates preserves the condition $N \cdot P = 0$ and the group law in the finite set E_N of N-torsion points which is rational over $\mathbb{Q}(j)$. This shows that $F_N(j)$ is a finite Galois extension of $\mathbb{Q}(j)$ with Galois group G obtained by restrictions of the $\sigma \in \mathrm{Aut}_{\mathbb{Q}(j)}\mathbb{C}$ and only depending on the associated group automorphism α of $E_N \simeq (\mathbb{Z}/N\mathbb{Z})^2$. The knowledge of the action of α on the X coordinate determines α up to sign, since a group automorphism such that $\alpha(P) \in \{\pm P\}$ for all P is ± 1. Thus one gets an injection of G in $\mathrm{Aut}(E_N)/\{\pm 1\}$. $\square$

We then have the following result.

LEMMA 3.64. *The field $F_N(j)$ contains a primitive N-th root of* 1.

PROOF. Given $\alpha \in E_N(j)$, there is by Lemma 3.62 a nonzero elliptic function

$$R_\alpha \in \mathbb{C}(x,y)/\{y^2 = x^3 - 3\,j\,(j-1728)\,x + 2\,j\,(j-1728)^2\},$$

unique up to multiplication by $\mathbb{C}^*$, which has simple zeros on $\{z \in E\,,\ Nz = \alpha\}$ and simple poles on E_N. The quotient $R_\alpha(z+\beta)/R_\alpha(z)$ only depends on the projective class of R_α and is therefore uniquely determined by α, β in $E_N(j)$. Thus, by Galois theory it belongs to the field generated by j and the coordinates of α and β. In fact, if one changes $\alpha \mapsto -\alpha$ and $\beta \mapsto -\beta$, the corresponding root of unity, constructed as above, does not change, as one can see using the θ-functions description. Thus it is fixed by any automorphism $\sigma \in \mathrm{Aut}_{F_N(j)}\mathbb{C}$ and belongs to $F_N(j)$. By Lemma 3.62 one can choose α and β so that the above quotient is a primitive N-th root of 1. Thus the field $F_N(j)$ contains a primitive N-th root of 1. $\square$

To summarize one gets (cf. [**268**] Theorem 6.6):

THEOREM 3.65. *The field F_N has the following properties.*

(1) *F_N is a Galois extension of $\mathbb{Q}(j)$ with Galois group*

$$\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$$

acting on the $\mathfrak{f}_u$ by permutation of the labels.

(2) *F_N contains $e^{2\pi i/N} \in \mathbb{C}$.*

(3) *$\mathbb{Q}(e^{2\pi i/N})$ is algebraically closed in F_N.*

(4) *Let $\tau \in \mathbb{H}$ be such that $j(\tau) \notin \bar{\mathbb{Q}}$. Then there exists an isomorphism ε_τ of F_N with $F_N(j) \subset \mathbb{C}$ uniquely determined by $\varepsilon_\tau(j) = j(\tau)$ and $\varepsilon_\tau(\mathfrak{f}_u) = \mathfrak{f}_u(\tau)$.*

PROOF. (1) First, notice that F_N is a countable field of meromorphic functions in $\mathbb{H}$. Thus, we can find a point τ_0 so that (4) holds for $\tau = \tau_0$. Thus, Lemma 3.63 implies that F_N is a finite Galois extension of $\mathbb{Q}(j)$ with Galois group given by $G \subset \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ acting explicitly by permutations of the labels of $\mathfrak{f}_u$. By construction $H_N \subset G$ and the fixed field $F_N^{H_N}$ contains the field $\mathbb{Q}(j, e^{2\pi i/N})$. It follows that $G = \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ since the product of the degrees of the extensions $F_N/F_N^{H_N}$ and $\mathbb{Q}(j, e^{2\pi i/N})/\mathbb{Q}(j)$ is the order of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$. It follows that $F_N^{H_N} = \mathbb{Q}(j, e^{2\pi i/N})$.
(2) and (3) follow immediately.

(4) For fixed $(\alpha, \beta, \gamma) \in \mathbb{Q}^{*3}$ with $\mathrm{Card}\{\alpha, \beta, \gamma\} = 3$ and for $N > 2$, the element

$$\xi = \alpha\, \mathfrak{f}_{(\frac{1}{N},0)} + \beta\, \mathfrak{f}_{(0,\frac{1}{N})} + \gamma\, \mathfrak{f}_{(\frac{1}{N},\frac{1}{N})} \tag{3.282}$$

is a generator of F_N over $\mathbb{Q}(j)$. Indeed, using the first terms in the q-expansion of the $\mathfrak{f}_u$ one checks that the subgroup of the Galois group that fixes ξ is trivial, hence ξ is a generator. One can then write the symmetric functions of ξ and the corresponding equation

$$P_{(\alpha,\beta,\gamma)}(\xi, j) = 0\,. \tag{3.283}$$

with coefficients in $\mathbb{Q}(j)$. Its degree is the order of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$.
If $j(\tau) \notin \bar{\mathbb{Q}}$, then the map $R(j) \in \mathbb{Q}(j) \to R(j(\tau)) \in \mathbb{C}$ is an isomorphism of fields with image a subfield of $\mathbb{C}$. The convergence of the series for $\mathfrak{f}_u(q)$ shows that $\xi(\tau)$ is a solution of (3.283) in $\mathbb{C}$ for $j = j(\tau)$. Thus, we get a morphism

$$F_N \to F_N(j(\tau)) \tag{3.284}$$

uniquely determined by the conditions $\varepsilon_\tau(j) = j(\tau)$ and $\varepsilon_\tau(\mathfrak{f}_u) = \mathfrak{f}_u(q)$, $q = e^{2\pi i\tau}$. ☐

We finally recall (cf. [**268**] Proposition 6.9) the characterization of elements of F_N among modular functions of level N in terms of the coefficients of their q-expansion.

Theorem 3.66. *F_N is the subfield of $F_N(\mathbb{C})$ of all elements whose expansion in powers of $q^{1/N}$ has all its coefficients in the cyclotomic field $\mathbb{Q}(e^{2\pi i/N})$.*

Proof. Let F'_N denote the subfield of $F_N(\mathbb{C})$ of all elements whose expansion in powers of $q^{1/N}$ has all its coefficients in the cyclotomic field $\mathbb{Q}(e^{2\pi i/N})$. It is easy to see that F'_N is a subfield, since the inverse is computed in simple rational terms. However, it has no reason *a priori* to be invariant (globally) under the action of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ on modular functions, since it is defined using a specific choice of a cusp. By construction and Proposition 3.59 one has $F_N \subset F'_N$. We need to show the equality. The field $F_N \vee \mathbb{C}$ generated by F_N and $\mathbb{C}$ is equal to $F_N(\mathbb{C})$, since the latter is a Galois extension of $\mathbb{C}(j) = \mathbb{Q}(j) \vee \mathbb{C}$ and no element of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ fixes all the $\mathfrak{f}_u$. Thus, we can write any $\rho \in F'_N$ in the form

$$\rho = \left(\sum \lambda_r\, x_r\right) / \left(\sum \mu_r\, y_r\right), \quad \text{with} \quad \lambda_r, \mu_r \in \mathbb{C} \quad \text{and} \quad x_r, y_r \in F_N. \tag{3.285}$$

This gives $\sum \mu_r y_r \rho - \sum \lambda_r x_r = 0$.
We can consider both $\mathbb{C}$ and F_N or F'_N as vector spaces over $\mathbb{Q}(e^{2\pi i/N})$. We need to show the injectivity of the natural morphism

$$\mathbb{C} \otimes_{\mathbb{Q}(e^{2\pi i/N})} F'_N \xrightarrow{\mu} F_N(\mathbb{C}) \tag{3.286}$$

given by $\mu(\lambda \otimes \rho) = \lambda\rho$. Suppose given $\lambda_i \in \mathbb{C}$, linearly independent over $\mathbb{Q}(e^{2\pi i/N})$, and $\rho_i \in F'_N$. Then the equality $\sum \lambda_i\, \rho_i = 0$ implies that the coefficient of $q^{k/N}$ in the q-expansion is 0 for all k. The coefficient of $q^{k/N}$ for ρ_i is in $\mathbb{Q}(e^{2\pi i/N})$. Moreover, $\sum \lambda_i\, c_i = 0$ with $c_i \in \mathbb{Q}(e^{2\pi i/N})$ implies $c_i = 0$ for all i. Thus, all coefficients of the q-expansion of ρ_i vanish, so that $\rho_i \equiv 0$.
Now, in the relation $\sum \mu_r y_r \rho = \sum \lambda_i x_i$ we can express μ_r and λ_i in terms of a basis (ξ_r) of $\mathbb{C}$ over $\mathbb{Q}(e^{2\pi i/N})$ and get an equality of the form $\sum \xi_r z_r \rho = \sum \xi_r t_r$ with

$z_r, t_r \in F_N$. It follows that $z_r\rho = t_r$ and both are non-zero for some r. This gives $\rho = t/z \in F_N$. □

6.3. Modular functions and modular forms.

Let $\mathcal{A}_N$ denote the subring of F_N of functions that are regular in $\mathbb{H}$, i.e. whose only poles are at the cusps. By construction $\mathcal{A}_N$ contains all polynomials in j as well as the $\mathfrak{f}_u$, but these do not suffice to generate it as a ring. A more precise description of the elements of $\mathcal{A}_N$ is given by the following result.

THEOREM 3.67. *Let $\mathcal{A}_N$ be the ring defined above.*

(1) *Let ω be a modular form of level N and weight $12k$, whose q-expansion has all of its coefficients in $\mathbb{Q}(e^{2\pi i/N})$. Then the function $f = \omega/\Delta^k$ is an element of $\mathcal{A}_N$.*
(2) *Every element of $\mathcal{A}_N$ is of the form described in (1) for a suitable k. One obtains in this way a filtration of $\mathcal{A}_N$ by weight.*

PROOF. (1) By construction, f is a modular function of level N. By Theorem 3.66 it belongs to F_N. Since ω is regular on H and Δ has no zero on $\mathbb{H}$, one gets that f is regular on $\mathbb{H}$, hence $f \in \mathcal{A}_N$.
(2) Suppose given $f \in \mathcal{A}_N$. Let q^{-k} be the largest negative power of q arising in the q-expansion of f at all cusps, $k \in \frac{1}{N}\mathbb{N}$. Let $k' \geq k$, $k' \in \mathbb{N}$. Then $\omega = \Delta^{k'} f$ is a modular form of level N, since no pole part remains at any of the cusps, and the $\Gamma(N)$ invariance follows from that of f. Moreover, all coefficients of the q-expansion of ω are in $\mathbb{Q}(e^{2\pi i/N})$ as required. The filtration is in fact that by the maximal order of poles at the cusps. □

COROLLARY 3.68. *Let ω be a modular form of level N whose q-expansion has all of its coefficients in $\mathbb{Q}(e^{2\pi i/N})$ for some cusp. Then the same holds at all cusps.*

PROOF. Let ω be of weight $2k$. Let $n, m \in \mathbb{N} \cup \{0\}$ with $2n + 3m + k = 0$ mod 6. Then, upon replacing ω by $e_4^n\, e_6^m\, \omega$, we can assume that $6|k$, and $12|2k$. Then $f = \omega\, \Delta^{-k/6} \in F_N$ and so does $f \mid \gamma$, for $\gamma \in \Gamma$. This gives the result. □

For each N, let $\mathcal{M}^{\mathrm{ab}}(N)$ denote the ring of modular forms of level N whose q-expansion has all of its coefficients in $\mathbb{Q}(e^{2\pi i/N})$. Let

$$\mathcal{M}^{\mathrm{ab}} = \bigcup_N \mathcal{M}^{\mathrm{ab}}(N). \tag{3.287}$$

Let η be the Dedekind η-function satisfying $\eta^{24} = \Delta$. This defines a map from $\mathcal{M}^{\mathrm{ab}}$ to modular functions by

$$\eta\left(\sum \omega_{2k}\right) = \sum \omega_{2k}\, \eta^{-4k}, \tag{3.288}$$

where ω_{2k} has weight $2k$.

PROPOSITION 3.69. *The map η of* (3.288) *is a surjective homomorphism from $\mathcal{M}^{\mathrm{ab}}$ to the algebra $\mathcal{O}_F = \cup_N \mathcal{A}_N$.*

PROOF. Notice that η^4 is of level 6. By construction the map η of (3.288) is a homomorphism and is surjective by (2) of Theorem 3.67. □

We now describe the kernel of the map η of (3.288).

LEMMA 3.70. *The kernel of the map* η *of* (3.288) *is the ideal generated by* $1-\eta^4$.

PROOF. Let J_1 be the kernel of the map η of (3.288) and let J_2 the ideal generated by $1-\eta^4$. One has $J_2 \subset J_1$. Moreover, modulo J_2 one has $\omega \sim \omega\, \eta^{4k}$. Thus, if $\omega = \sum \omega_{2k} \in J_1$, then $\sum \eta^{4(n-k)}\, \omega_{2k} = 0$ and $\sum \omega_{2k} \sim 0$ modulo J_2 so that $J_1 = J_2$. $\square$

REMARK 3.71. *Notice that the ideal* $J = \operatorname{Ker} \eta$ *is stable under the Ramanujan derivation* $\mathcal{X}$*:*

$$\mathcal{X} := \frac{1}{2\pi i}\frac{d}{dz} - \frac{1}{12\pi i}\,\frac{d}{dz}(\log \Delta)\cdot Y = \frac{1}{2\pi i}\,\frac{d}{dz} - \frac{1}{2\pi i}\,\frac{d}{dz}(\log \eta^4)\cdot Y\,, \tag{3.289}$$

where Y *stands for the grading operator*

$$Y(f) = kf\,, \quad \forall\, f \in \mathcal{M}_{2k}^{\mathrm{ab}}\,. \tag{3.290}$$

Thus, $\mathcal{X}$ *descends to the algebra* $\mathcal{O}_F$*. In fact, it is simply the derivation*

$$\mathcal{X}(f) = \frac{1}{2\pi i}\,\frac{df}{\eta^4 dz}\,, \quad \forall f \in \mathcal{O}_F\,, \tag{3.291}$$

where $\eta^4 dz$ *is an invariant form under* $\Gamma(6)$.

In the level 1 case Proposition 3.57 gives a simple characterization of the elements of $\mathcal{A}_1 \subset F_1$ as polynomials in j with rational coefficients.

PROPOSITION 3.72. *Let* $f \in F_N$*. Then* $f \in \mathcal{A}_N$ *if and only if* f *is integral over* $\mathcal{A}_1$*, i.e. if it satisfies an equation of the form*

$$f^n + \sum_{k=1}^{n} P_k(j) f^{n-k} = 0, \qquad \text{for } P_k \in \mathbb{Q}[X], \tag{3.292}$$

PROOF. If f satisfies (3.292) then $f \in \mathcal{A}_N$. Indeed, for $\tau \in \mathbb{H}$, let v_τ be the valuation on F_N which assigns to $f \in F_N$ its order at τ. The latter is equal to $-k$ if f has a pole of order k. This satisfies

$$v_\tau(f+g) \geq \inf\{v_\tau(f), v_\tau(g)\}, \qquad \text{and} \qquad v_\tau(fg) = v_\tau(f) + v_\tau(g). \tag{3.293}$$

so that for $v_\tau(f) = \alpha < 0$ one gets $n\alpha \geq \inf\{(n-1)\alpha, \ldots, \alpha, 0\}$, which gives a contradiction.

Conversely, let $f \in \mathcal{A}_N$. We first view it as an element of $\mathcal{A}_N(\mathbb{C})$, i.e. as a modular function of level N. For all $\gamma \in \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ one has $f \mid \gamma \in \mathcal{A}_N$. Then all the symmetric functions s_k of the $f \mid \gamma$ belong to the intersection of $F_N^{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}$ with $\mathcal{A}_N$, which is the ring of polynomials $\mathbb{Q}[j, e^{2\pi i/N}]$. Indeed, the fixed field is $\mathbb{Q}(j)[e^{2\pi i/N}]$ and the regularity on $\mathbb{H}$ then implies that these elements are actually in the ring $\mathbb{Q}[j, e^{2\pi i/N}]$. Thus one obtains the answer using the Galois group of the cyclotomic extension. $\square$

6.4. Explicit computations for $N = 2$ and $N = 4$.

The field F_N for $N = 2$ is generated by the division values at points of order 2. In terms of the elliptic curve (3.273) with coordinates X and Y, this corresponds to points with $-P = P$, i.e. with $Y = 0$. This gives the three zeros of the resulting polynomial in X.

Let $a = \wp_\Lambda\left(\frac{1}{2}\right)$, $b = \wp_\Lambda\left(\frac{\tau}{2}\right)$ and $c = \wp_\Lambda\left(\frac{1+\tau}{2}\right)$, with $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$. The homogeneous functions of degree zero in a, b, c generate $F_{N=2}$. In fact, one has

$$\lambda = \frac{c-b}{a-b} = k^2 = \frac{\theta_2(0)^4}{\theta_3(0)^4}, \tag{3.294}$$

where the θ-functions are taken relative to the modulus $q^{1/2}$. This means that, for instance, one has

$$\theta_3(z) = \sum_{n\in\mathbb{Z}} q^{\frac{n^2}{2}} e^{2\pi\, inz}. \tag{3.295}$$

One then gets $F_{N=2} = \mathbb{Q}(\lambda)$, the field of rational fractions in λ. The subfield $\mathbb{Q}(j)$ is generated by

$$j = 256\frac{(1-\lambda+\lambda^2)^3}{\lambda^2(1-\lambda)^2}. \tag{3.296}$$

The Galois group of $F_{N=2}$ over $\mathbb{Q}(j)$ is geometric, given by $\mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$ acting as the dihedral group S_3 of order 6, generated by the transformations $\lambda \mapsto 1-\lambda$ and $\lambda \mapsto 1/\lambda$. It corresponds to the ambiguity in the cross ratio of four elements (a, b, c, d) of $\mathbb{P}^1(\mathbb{C})$. Once this ambiguity is lifted, one gets, instead of (3.273), the equation

$$y^2 = (1-x^2)(1-k^2\, x^2), \quad \text{with } k^2 = \lambda, \tag{3.297}$$

which is parameterized by $x = sn(u)$ and $y = cn(u)\, dn(u)$, in terms of the Jacobi elliptic functions

$$sn(u) = k^{-1/2}\frac{\theta_1(u)}{\theta_4(u)}, \quad cn(u) = k^{-1/2}k'^{1/2}\frac{\theta_2(u)}{\theta_4(u)}, \quad \text{and} \quad dn(u) = k'^{1/2}\frac{\theta_3(u)}{\theta_4(u)},$$

where $k^{1/2} = \frac{\theta_2(0)}{\theta_3(0)}$ and $k'^{1/2} = \frac{\theta_4(0)}{\theta_3(0)}$. This leads us to the case of level $N = 4$.

In the discussion above both k and k' are of level 4 and they fulfill the equation $k^2 + k'^2 = 1$. This is better parameterized by t with

$$t = \frac{k'}{1+k}, \quad \text{so that } k = \frac{1-t^2}{1+t^2}, \quad \text{and} \quad k' = \frac{2t}{1+t^2}.$$

One gets $F_{N=4} = \mathbb{Q}(t, i)$. The Galois group of $F_{N=4}$ over $\mathbb{Q}(j)$ is

$$\mathrm{Gal}(F_{N=4}/\mathbb{Q}(j)) = \mathrm{GL}_2(\mathbb{Z}/4\mathbb{Z})/\{\pm 1\},$$

which has order 48. It contains the normal subgroup $\mathrm{SL}_2(\mathbb{Z}/4\mathbb{Z})/\{\pm 1\} \simeq S_4$, which acts geometrically, with the Klein subgroup acting by altering the signs $k \mapsto \pm k$ and $k' \mapsto \pm k'$. The division values $\mathfrak{f}_u$ of order four are given by explicit rational functions of t, but with coefficients in $\mathbb{Q}(i)$. The remaining part of the Galois group action is the cyclotomic action on the coefficients through the determinant.

6.5. The modular field F and $\mathbb{Q}$-lattices.

The modular field F is defined as

$$F = \cup_N F_N. \tag{3.298}$$

It is by construction a *countable* field of meromorphic functions of $\tau \in \mathbb{H}^*$. We show in the next result how to describe all embeddings $F \overset{\iota}{\hookrightarrow} \mathbb{C}$ in terms of invertible $\mathbb{Q}$-lattices.

PROPOSITION 3.73. *Let $F = \cup_N F_N$ be the modular field defined above.*

(1) *Let $L = (\Lambda, \varphi)$ be an* invertible *$\mathbb{Q}$-lattice such that $j(\Lambda) \notin \bar{\mathbb{Q}}$. Then there exists a unique embedding $F \overset{\iota_L}{\hookrightarrow} \mathbb{C}$, such that*

$$\iota_L(\mathfrak{f}_a) = -2^7\, 5^3\, \frac{g_2\, g_3}{\Delta}\, (\Lambda) \wp_\Lambda(\varphi(a)), \tag{3.299}$$

for all $a \in \mathbb{Q}^2/\mathbb{Z}^2$, $a \neq 0$.

(2) *All embeddings $\iota : F \hookrightarrow \mathbb{C}$ are of the form* (3.299).

PROOF. (1) The $\mathfrak{f}_a$, for $a \in \mathbb{Q}^2/\mathbb{Z}^2$, generate F, since j is a function of the 2-torsion values. Thus, the uniqueness follows. To get the existence one can assume that Λ is of the form $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$ and $\varphi(a) = \rho_1(a) - \tau\rho_2(a)$, for some $\rho \in \mathrm{GL}_2(\hat{\mathbb{Z}}) \subset M_2(\hat{\mathbb{Z}})$. Since $j(\tau) = j(\Lambda) \notin \bar{\mathbb{Q}}$, the evaluation at τ defines an embedding

$$I_\tau : F \hookrightarrow \mathbb{C}. \tag{3.300}$$

On the other hand, the Galois group of $F/\mathbb{Q}(j)$ is the compact group

$$\mathrm{GL}_2(\hat{\mathbb{Z}})/\{\pm 1\} = \varprojlim_N \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}\,,$$

acting by permutations of the labels. Upon replacing I_τ by $I_\tau \circ \alpha$ where α is the Galois automorphism of $F/\mathbb{Q}(j)$ such that $\alpha(\mathfrak{f}_a) = \mathfrak{f}_{\rho(a)}$ for all a one gets the required embedding.

(2) Let $\iota : F \hookrightarrow \mathbb{C}$. Then since the restriction of ι to $\mathbb{Q}(j)$ is injective one gets $\iota(j) \notin \bar{\mathbb{Q}}$. Consider a point $\tau \in \mathbb{H}$ such that $\iota(j) = j(\tau)$. Then $\iota(F) = I_\tau(F)$ as subfields of $\mathbb{C}$, since both are obtained as the fields of X-coordinates of torsion points in the elliptic curve $E(j)$ given by (3.273), $E(j) \sim \mathbb{C}/\Lambda$ with $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$. Then $I_\tau^{-1} \circ \iota \in \mathrm{Aut}(F/\mathbb{Q}(j))$, and this determines $\rho \in \mathrm{GL}_2(\hat{\mathbb{Z}})/\{\pm 1\}$, as well as the required invertible $\mathbb{Q}$-lattice (Λ, φ). □

LEMMA 3.74. (1) *As a field of meromorphic functions on $\mathbb{H}$, F is stable under $f \mapsto f \mid \gamma$ for any $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$.*

(2) *Let $L = (\Lambda, \varphi)$ and $L' = (\Lambda', \varphi')$ be* invertible *$\mathbb{Q}$-lattices that are in the same orbit for the right action of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ on the quotient $\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}})$. Then, if $j(\Lambda) \notin \bar{\mathbb{Q}}$,*

$$\iota_L(F) = \iota_{L'}(F) \subset \mathbb{C}.$$

PROOF. (1) This is clear for $\gamma \in \Gamma = \mathrm{SL}_2(\mathbb{Z})$. Also, by looking at the q-expansion for $\Im\tau \to \infty$, one sees that it holds for $\gamma(\tau) = \frac{a\tau+b}{d}$. Thus, it holds in general since these two subgroups generate $\mathrm{GL}_2(\mathbb{Q})$.

(2) The range $\iota_L(F)$ is independent of the labeling φ of the torsion points and only depends upon the lattice Λ. The hypothesis implies that the two lattices Λ and Λ' are commensurable. With $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$ one can choose for Λ' a basis $(c\tau + d, a\tau + b)$ with $a, b, c, d \in \mathbb{Q}$ and

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{Q}).$$

Let $\tau' = \frac{a\tau+b}{c\tau+d}$. Then $I_{\tau'}(F) = I_\tau(F)$, using (1), so that $\iota_{L'}(F) = \iota_L(F)$. □

We then have the following result of Shimura [**268**] with slightly different notation.

THEOREM 3.75. *There exists a homomorphism $\alpha \mapsto \mathrm{Gal}(\alpha)$ of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ to $\mathrm{Aut}_{\mathbb{Q}}(F)$ such that*

a) *For* $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$*, one has* $\mathrm{Gal}(\alpha)(f) = f \mid \alpha^{-1}$*, for all* $f \in F$*.*
b) *For* $\alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})$*, one has* $\mathrm{Gal}(\alpha)\mathfrak{f}_a = \mathfrak{f}_{\alpha(a)}$*, for all* $a \in \mathbb{Q}^2/\mathbb{Z}^2$*,* $a \neq 0$*.*

PROOF. We view the label u as a column vector and act on it by $\alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ accordingly. Notice that our notation for the $\mathfrak{f}_u$ differs from the standard one $\mathbf{f}_a$ used in Shimura [**268**]. They are related by

$$\mathfrak{f}_u = \mathbf{f}_{(su)^t}, \quad \text{with } s = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.301}$$

Indeed in [**268**] one uses the row vector $u^t = (a_1, a_2)$ and the corresponding torsion point is

$$u^t \cdot \begin{pmatrix} \tau \\ 1 \end{pmatrix} = a_1\tau + a_2 = \lambda \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix} = \lambda(su) \tag{3.302}$$

By [**268**] there exists an antihomomorphism τ from $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ to the Galois group $\mathrm{Aut}_{\mathbb{Q}}(F)$ such that the following basic formulas hold ([**268**], cf. 6.2.1 p.147 and 6.6.1 p.149):

$$\mathbf{f}_a^{\tau(u)} = \mathbf{f}_{au}, \quad \forall u \in \mathrm{GL}_2(\hat{\mathbb{Z}}), \qquad \text{and } f^{\tau(\alpha)} = f \mid \alpha, \quad \forall \alpha \in \mathrm{GL}_2^+(\mathbb{Q}). \tag{3.303}$$

Here $\tau(g_1\, g_2) = \tau(g_2)\,\tau(g_1)$. Thus, we prefer to compose τ with the antihomomorphism from $\mathrm{GL}_2(A_f)$ to itself given by

$$\bar{\alpha} = \alpha^{-1} \det(\alpha). \tag{3.304}$$

One then obtains

$$\bar{\alpha} = \begin{pmatrix} \alpha_{22} & -\alpha_{12} \\ -\alpha_{21} & \alpha_{11} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = s\alpha^t s^{-1}.$$

We then set

$$\mathrm{Gal}(\alpha) = \tau(\bar{\alpha}). \tag{3.305}$$

For $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$, one has

$$\mathrm{Gal}(\alpha) f = f^{\tau(\bar{\alpha})} = f \mid \tau(\bar{\alpha}),$$

but the scalar $\det\alpha$ plays no role in this composition and one gets $f \mid \alpha^{-1}$ as in (a). For $\alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})$, one obtains

$$\mathbf{f}_{(sa)^t}^{\tau(\bar{\alpha})} = \mathbf{f}_{(sa)^t\bar{\alpha}}.$$

Since $(sa)^t = a^t\, s^t$ and $\bar{\alpha} = s\,\alpha^t\, s^{-1}$, while signs do not matter, one has $(sa)^t\bar{\alpha} = a^t\, s^t s\,\alpha^t\, s^{-1}$ and obtains $\mathbf{f}_{(sa)^t\bar{\alpha}} = \mathbf{f}_{(s\alpha a)^t}$, which corresponds to $a \mapsto \alpha a$. $\square$

In order to get a better formulation of the case (b), we first consider the special case of elements of $\mathrm{GL}_2(\hat{\mathbb{Z}})$ that are of the form

$$\alpha_v = \begin{pmatrix} v & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{for} \quad v \in \mathrm{GL}_1(\hat{\mathbb{Z}}). \tag{3.306}$$

For these elements we have the following result.

LEMMA 3.76. *Let* $f = \sum_n a_n\, q^{n/N} \in F_N$ *and let* v *and* α_v *be as in* (3.306)*. Then the series* $\sum_n \mathrm{Gal}(v)(a_n) q^{n/N}$ *defines an element in* F_N *which is equal to* $\mathrm{Gal}(\alpha_v)\, f$*.*

PROOF. Here we let $\mathrm{GL}_1(\hat{\mathbb{Z}})$ act on the cyclotomic field $\mathbb{Q}^{\mathrm{ab}}$ as the Galois group $\mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q})$. Thus, with $p_N(v) = v_N \in (\mathbb{Z}/N\mathbb{Z})^*$, one gets

$$\mathrm{Gal}(v)\, e^{2\pi i/N} = e^{2\pi i v_N/N}. \tag{3.307}$$

The map $\sum_n a_n q^{n/N} \mapsto \sum_n \mathrm{Gal}(v)(a_n)\, q^{n/N}$ is an automorphism of the field of formal series with coefficients in $\mathbb{Q}^{\mathrm{ab}}$, which is the identity on $\mathbb{Q}(j)$. The map $\mathrm{Gal}(\alpha_v)$ is an automorphism of F_N over $\mathbb{Q}(j)$, and F_N is generated by the $\mathfrak{f}_u$. Thus, it is enough to check that those two maps agree on the generators $\mathfrak{f}_u$. Thus, we can start with

$$\mathfrak{f}_u = \left(-2^7\, 3^5\, \frac{g_2\, g_3}{\Delta}\right) \left(\frac{1}{12} + \frac{z}{(1-z)^2} + \sum_n \sum_{d|n} d(z^d + z^{-d} - 2)q^n\right) \tag{3.308}$$

and $z = e^{2\pi i a/N}\, q^{-b/N}$.
The first map acts on $\mathfrak{f}_u$ by replacing $z = e^{2\pi i a/N} q^{-b/N}$ with

$$z' = \mathrm{Gal}(v)(e^{2\pi i a/N})\, q^{-b/N} = e^{2\pi i a v_N/N}\, q^{-b/N}.$$

This amounts to replacing a by av_N and gives $\mathfrak{f}_{\alpha_v(u)}$, which is the action of $\mathrm{Gal}(\alpha_v)$ on the generator $\mathfrak{f}_u$. ☐

Theorem 3.75 is due to [**268**]. It is worthwhile to give the explicit construction in the above notations. We let $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})^0$ denote the subgroup of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ specified by the condition

$$\det \rho \in \mathrm{GL}_1^+(\mathbb{Q}) \subset \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},f}). \tag{3.309}$$

The density of $\mathrm{SL}_2(\mathbb{Q})$ in $\mathrm{SL}_2(\mathbb{A}_{\mathbb{Q},f})$ implies the density of $\mathrm{GL}_2^+(\mathbb{Q})$ inside $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})^0$. This gives an extension by continuity of the action $f \mapsto f \mid \gamma^{-1}$ of $\mathrm{GL}_2^+(\mathbb{Q})$ on F to an action Gal_1 of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})^0$, since the isotropy group of a given $f \in F$ contains the open subgroup of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})^0$ associated to the closure $\overline{\Gamma(N)}$ of $\Gamma(N)$.
Now a given $g \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ can be written as $g = \rho\alpha_u = \alpha_u\rho'$, where $u \in \mathrm{GL}_1(\hat{\mathbb{Z}})$ gives the $\mathrm{GL}_1(\hat{\mathbb{Z}})$ component of $\det g$ and $\rho, \rho' \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})^0$. One needs to show that, with the above notation, we have

$$\mathrm{Gal}_1(\rho)\,\mathrm{Gal}(\alpha_u) = \mathrm{Gal}(\alpha_u)\mathrm{Gal}_1(\rho') \tag{3.310}$$

where Gal_1 is the natural action of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})^0$.
Let us first take $\rho = \gamma\,\delta\,\gamma'$, where $\gamma, \gamma' \in \mathrm{SL}_2(\hat{\mathbb{Z}})$ and δ is diagonal. The effect of $\mathrm{Gal}_1(\delta)$ on the q-series of f is to replace q by q^r, where r is a rational number. This operation commutes with the action of $\mathrm{Gal}(\alpha_u)$, so that $\mathrm{Gal}_1(\delta)\mathrm{Gal}(\alpha_u) = \mathrm{Gal}(\alpha_u)\mathrm{Gal}_1(\delta)$. We let $\gamma_u = \alpha_u^{-1}\gamma\alpha_u$ and $\gamma'_u = \alpha_u^{-1}\gamma'\alpha_u$. We also have $\mathrm{Gal}_1(\gamma)\mathrm{Gal}(\alpha_u) = \mathrm{Gal}(\alpha_u)\mathrm{Gal}_1(\gamma_u)$, which follows from the form of the Galois action of $\mathrm{GL}_2(\hat{\mathbb{Z}})$ on F by permutations of the $\mathfrak{f}_u$. It then follows that

$$\begin{aligned}
\mathrm{Gal}_1(\rho)\mathrm{Gal}(\alpha_u) &= \mathrm{Gal}_1(\gamma)\mathrm{Gal}_1(\delta)\mathrm{Gal}_1(\gamma')\mathrm{Gal}(\alpha_u)\\
&= \mathrm{Gal}_1(\gamma)\mathrm{Gal}_1(\delta)\mathrm{Gal}(\alpha_u)\mathrm{Gal}_1(\gamma'_u)\\
&= \mathrm{Gal}_1(\gamma)\mathrm{Gal}(\alpha_u)\mathrm{Gal}_1(\delta)\mathrm{Gal}_1(\gamma'_u)\\
&= \mathrm{Gal}(\alpha_u)\mathrm{Gal}_1(\gamma_u)\mathrm{Gal}_1(\delta)\mathrm{Gal}_1(\gamma'_u)\\
&= \mathrm{Gal}(\alpha_u)\mathrm{Gal}_1(\rho')
\end{aligned}$$

since $\rho' = \alpha_u^{-1}\rho\alpha_u = \gamma_u\delta\gamma'_u$. The general case then follows. ☐

7. Arithmetic of the GL_2 system

We can now describe the arithmetic properties of the quantum statistical mechanical system $(\mathcal{A}_2, \sigma_t)$. A fundamental step is the construction of an arithmetic algebra $\mathcal{A}_{2,\mathbb{Q}}$ that will play a role analogous to the arithmetic subalgebra $\mathcal{A}_{1,\mathbb{Q}}$ of the Bost–Connes system described in Section 4.4. In the GL(2) case, the algebra $\mathcal{A}_{2,\mathbb{Q}}$ will not be a subalgebra of $\mathcal{A}_2$ but a suitable subalgebra of an algebra of unbounded multipliers of $\mathcal{A}_2$. We will see that the extremal KMS_∞ states of $(\mathcal{A}_2, \sigma_t)$ extend to this algebra $\mathcal{A}_{2,\mathbb{Q}}$ of unbounded multipliers and, for a *generic* choice of an extremal state $\varphi \in \mathcal{E}_\infty$, the values $\varphi(\mathcal{A}_{2,\mathbb{Q}})$ generate a subfield $F_\varphi \subset \mathbb{C}$ isomorphic to the modular field F.

7.1. The arithmetic subalgebra: explicit elements.

To construct the arithmetic algebra $\mathcal{A}_{2,\mathbb{Q}}$ we use the well known analogy between trigonometric functions and elliptic functions explained in [**297**]. We view the description of §4.4 of the arithmetic subalgebra $\mathcal{A}_{1,\mathbb{Q}}$ of the Bost–Connes system as the trigonometric case and pass to the elliptic case. This gives natural candidates for elements of $\mathcal{A}_{2,\mathbb{Q}}$ which we describe in this section. In the next section §7.2 we shall abstract their properties and define $\mathcal{A}_{2,\mathbb{Q}}$ implicitly from algebraic conditions on its elements.

As in the one-dimensional case, we first deal with functions $f(g, \rho, \tau)$ that vanish except for $g \in \Gamma$. By construction these are functions on the space

$$\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H}) \simeq (\text{Space of } \mathbb{Q}\text{-lattices in } \mathbb{C})/\mathbb{C}^* \tag{3.311}$$

of 2-dimensional $\mathbb{Q}$-lattices up to scaling.

These functions can be considered as functions of weight zero using the following analogue of Definition 3.26.

Definition 3.77. *A function f on the space of 2-dimensional $\mathbb{Q}$-lattices is homogeneous of weight k if it satisfies*

$$f(\lambda(\Lambda, \phi)) = \lambda^{-k} f(\Lambda, \phi), \qquad \forall \lambda \in \mathbb{C}^*. \tag{3.312}$$

To obtain explicit elements of $\mathcal{A}_{2,\mathbb{Q}}$ we start with Eisenstein series and view them as functions on the space of 2-dimensional $\mathbb{Q}$-lattices. They are labeled by $a \in (\mathbb{Q}/\mathbb{Z})^2$ and given for $k > 1$ by

$$E_{2k,a}(\Lambda, \phi) = (2\pi i)^{-2k} \sum_{y \in \Lambda + \phi(a)} y^{-2k}. \tag{3.313}$$

As in the one-dimensional case this is undefined when $\phi(a) \in \Lambda$. We show below how to deal with this problem (cf. (3.319)). By construction $E_{2k,a}$ is homogeneous of weight $2k$. Recall that to a pair $(\rho, \tau) \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ we associate (cf. (3.165)) a $\mathbb{Q}$-lattice given by

$$L(\rho, \tau) = (\Lambda, \phi)\,, \quad \Lambda = \mathbb{Z} + \tau\mathbb{Z}\,,\ \phi(a) = \rho_1(a) - \tau\rho_2(a) \in \mathbb{Q}\Lambda/\Lambda, \tag{3.314}$$

where $\rho_j(a) = \sum \rho_{jk}(a_k) \in \mathbb{Q}/\mathbb{Z}$, for $a = (a_1, a_2) \in (\mathbb{Q}/\mathbb{Z})^2$. This allows us to view $E_{2k,a}$ as a function of (ρ, τ),

$$E_{2k,a}(\rho, \tau) = (2\pi i)^{-2k} \sum_{y \in \Lambda + \phi(a)} y^{-2k}\,, \quad (\Lambda, \phi) = L(\rho, \tau)\,. \tag{3.315}$$

For $\gamma \in \Gamma = \mathrm{SL}_2(\mathbb{Z})$ we have

$$E_{2k,a}(\gamma\rho, \gamma\tau) = (c\tau + d)^{2k} E_{2k,a}(\rho, \tau).$$

This follows from the homogeneity of weight $2k$ since one has

$$L(\gamma\rho, \gamma(\tau)) = (cz + d)^{-1}\, L(\rho, \tau)\,, \quad \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \tag{3.316}$$

as can be checked directly since

$$(a\rho_1 + b\rho_2)(c\tau + d) - (c\rho_1 + d\rho_2)(a\tau + b) = \rho_1 - \tau\rho_2\,.$$

The analogue of the function c of (3.79) is given by the function

$$\mathbf{c}(\tau) = -2^7\, 3^5\, \frac{g_2 g_3}{\Delta} \tag{3.317}$$

which has weight -2 and no pole in $\mathbb{H}$. Thus, the functions $\mathcal{E}_{2k,a} = \mathbf{c}^k\, E_{2k,a}$ are homogeneous of weight zero, and are natural arithmetic functions of $\mathbb{Q}$-lattices, but we still need to define their value when $\phi(a) \in \Lambda$. We proceed as in the one-dimensional case and first consider the limit case $k = 1$. For $k = 1$ we let, with $(\Lambda, \phi) = L(\rho, \tau)$,

$$X_a(\rho, \tau) = (2\pi i)^{-2} \left(\sum_{y \in \Lambda + \phi(a)} y^{-2} - \sum_{y \in \Lambda^*} y^{-2} \right). \tag{3.318}$$

This makes sense when $\phi(a) \notin \Lambda$ and is just the evaluation of the Weierstrass $\wp$-function on $\phi(a)$ as in (3.276). We let $X_a(\rho, \tau) = 0$ if $\phi(a) \in \Lambda$. In (3.318) the notation Λ^* stands for $\Lambda \smallsetminus \{0\}$ and the sum is convergent after grouping the terms corresponding to the same element of Λ^*. We let $\mathcal{X}_a := \mathbf{c}\, X_a$. It is a continuous function on $\Gamma \backslash (M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$. The Γ-invariance follows from (3.316) since the product $\mathbf{c}\, X_a$ is of weight zero. The continuity of $\mathcal{X}_a$ as a function of ρ comes from the fact that it only involves the restriction $\rho_N \in M_2(\mathbb{Z}/N\mathbb{Z})$ of ρ to N-torsion elements when the label a fulfills $Na = 0$.

PROPOSITION 3.78. *There exists a unique normalization*

$$E_{2k}(\rho, \tau) = \nu_{2k}(\tau) \quad \text{if} \quad \varphi(a) \in \Lambda, \tag{3.319}$$

such that the following relation holds for all m (with $E_{2,a} - e_2 := X_a$):

$$\begin{aligned} E_{2m,a} \;&=\; X_a(E_{2m-2,a} - e_{2m-2}) + \left(1 - \binom{2m}{2}\right) e_{2m} \\ &\quad - \sum_{k=1}^{m-2} \binom{2k+1}{2k} e_{2k+2}(E_{2m-2k-2,a} - e_{2m-2k-2})\,. \end{aligned} \tag{3.320}$$

Moreover ν_{2k} is a modular form of weight $2k$. One has $\nu_{2k} \in \mathbb{Q}[e_4, e_6]$ and the first values are

$$\nu_4 = -5\, e_4\,, \quad \nu_6 = -14\, e_6\,, \quad \nu_8 = \frac{45}{7}\, e_4^2\,, \ldots \tag{3.321}$$

PROOF. By equation (3) Chapter V of [**297**] with the pair (n, m) there given by $(2, 2m-2)$ one gets (using $e_k = 0$ for k odd) that (3.320) holds as long as $\varphi(a) \notin \Lambda$. Since we want the same relation to hold for $\varphi(a) \in \Lambda$ and have $X_a = 0$ in that case, we obtain an inductive relation on the $\nu_{2k}(\tau)$ of the form

$$\begin{aligned} \nu_{2m} = & \left(1 - \binom{2m}{2}\right) e_{2m} \\ - & \textstyle\sum_{k=1}^{m-2} \binom{2k+1}{2k} e_{2k+2}(\nu_{2m-2k-2} - e_{2m-2k-2}) \,. \end{aligned} \tag{3.322}$$

Using (3.322) one gets that ν_{2k} is a modular form of weight $2k$ which is a polynomial with rational coefficients in the e_{2n}. Moreover the e_{2n} for $n \geq 2$ are in the ring $\mathbb{Q}[e_4, e_6]$ (cf [**297**]) thanks to the relation

$$\frac{1}{3}(m-3)(4m^2-1)e_{2m} = \sum_{2}^{m-2} (2r-1)(2m-2r-1)e_{2r}e_{2m-2r}. \tag{3.323}$$

□

We now obtain the analogue of Lemma 3.27 in the two-dimensional case.

PROPOSITION 3.79. *Let $\mathcal{X}_a := \mathbf{c}\, X_a$ and $\mathcal{E}_{2k,a} = \mathbf{c}^k\, E_{2k,a}$. There exist polynomials P_n in two variables, with rational coefficients, such that*

$$\mathcal{E}_{2k,a} = P_k(\mathcal{X}_a, j). \tag{3.324}$$

The first ones are given by

$$\begin{aligned} &P_2 = X^2 - j(j-1728) \\ &P_3 = X^3 - \tfrac{9}{5}\, Xj(j-1728) + \tfrac{4}{5} j(j-1728)^2 \\ &\dots \end{aligned}$$

PROOF. Since (3.320) is homogeneous it can be written in the form

$$\begin{aligned} \mathcal{E}_{2m,a} \;=\; & \mathcal{X}_a(\mathcal{E}_{2m-2,a} - \mathbf{c}^{m-1}\, e_{2m-2}) + \left(1 - \binom{2m}{2}\right) \mathbf{c}^m\, e_{2m} \\ - & \sum_{k=1}^{m-2} \binom{2k+1}{2k} \mathbf{c}^{k+1}\, e_{2k+2}(\mathcal{E}_{2m-2k-2,a} - \mathbf{c}^{m-k-1}\, e_{2m-2k-2}) \,. \end{aligned} \tag{3.325}$$

Moreover using (3.323) one shows by induction that $\mathbf{c}^m\, e_{2m}$ is in the ring $\mathbb{Q}[\mathbf{c}^2\, e_4, \mathbf{c}^3\, e_6]$. A direct computation gives

$$\mathbf{c}^2\, e_4 = \frac{1}{5}\, j\,(j-1728), \qquad \text{and} \qquad \mathbf{c}^3\, e_6 = -\frac{2}{35}\, j\,(j-1728)^2.$$

Thus $\mathbf{c}^m\, e_{2m} \in \mathbb{Q}[j]$. The proof of the proposition then follows by induction on k. □

For the BC system Theorem 3.30 shows that the arithmetic subalgebra $\mathcal{A}_{1,\mathbb{Q}}$ is generated by the Eisenstein series $e_{1,a}$ and the μ_n, μ_n^*. The generators μ_n are essential in order to implement the commensurability relation of $\mathbb{Q}$-lattices.

In the case of two-dimensional $\mathbb{Q}$-lattices, the commensurability relation is implemented by the following arithmetic elements T_{r_1,r_2} for $r_j \in \mathbb{Q}_+^*$, with $r_1|r_2$, associated to the classical Hecke correspondences. We let $C_{r_1,r_2} \subset \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$ be the finite subset given by the double coset class of

$$\begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix} \in \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})/\Gamma.$$

We then define

$$T_{r_1,r_2}(g,\rho,\tau) = \begin{cases} 1 & g \in C_{r_1,r_2}, \text{ with } g\rho \in M_2(\hat{\mathbb{Z}}) \\ 0 & \text{otherwise.} \end{cases} \tag{3.326}$$

As will be shown below in Proposition 3.82 the T_{r_1,r_2} will define elements of the arithmetic subalgebra $\mathcal{A}_{2,\mathbb{Q}}$ and by Proposition 3.87 fulfill the relations of the Hecke algebra $\mathcal{H}(\mathrm{GL}_2^+(\mathbb{Q}),\Gamma)$ of double cosets, provided r_1, r_2 are restricted to vary among positive *integers*.

7.2. The arithmetic subalgebra: definition.

Let $\tilde{\mathcal{U}}_2 \subset \mathrm{GL}_2^+(\mathbb{Q}) \times M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ be as in (3.194) and $\mathcal{U}_2 = (\Gamma\times\Gamma)\backslash\tilde{\mathcal{U}}_2$ be its quotient by the action of $\Gamma\times\Gamma$ as in (3.195). Let $f \in C(\mathcal{U}_2)$ be a function with *finite support* in the variable $g \in \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$. Any such function f defines an unbounded multiplier on the C^*-algebra $\mathcal{A}_2$ using the convolution product (3.196).

The *arithmetic* algebra $\mathcal{A}_{2,\mathbb{Q}}$ will now be obtained as a subalgebra of this algebra of multipliers. The $\mathcal{X}_a$, $\mathcal{E}_{2m,a}$ and the T_{r_1,r_2} of §7.1 are the basic examples of *arithmetic* elements. By construction, the T_{r_1,r_2} are $\Gamma\times\Gamma$-invariant functions on $\tilde{\mathcal{U}}_2$. The $\mathcal{X}_a$ and $\mathcal{E}_{2m,a}$ are continuous functions on $\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})$ but we extend them to $\Gamma\times\Gamma$-invariant functions on $\tilde{\mathcal{U}}_2$ by

$$\mathcal{X}_a(\gamma,\rho,\tau) := \mathcal{X}_a(\rho,\tau), \quad \forall\gamma\in\Gamma\,, \quad \mathcal{X}_a(\gamma,\rho,\tau) = 0\,, \quad \forall\gamma\notin\Gamma \tag{3.327}$$

and similarly for the $\mathcal{E}_{2m,a}$.

Our first task is to abstract the arithmetic properties fulfilled by these specific functions. We first define the notion of *level* of a function $f \in C(\mathcal{U}_2)$. Let $p_N : M_2(\hat{\mathbb{Z}}) \to M_2(\mathbb{Z}/N\mathbb{Z})$ be the canonical projection. It is a ring homomorphism.

DEFINITION 3.80. *A function $f \in C(\mathcal{U}_2)$ has level N iff $f(g,\rho,z)$ only depends upon*

$$(g, p_N(\rho), z) \in \mathrm{GL}_2^+(\mathbb{Q}) \times M_2(\mathbb{Z}/N\mathbb{Z}) \times \mathbb{H}.$$

In this case, specifying f with finite support in the variable $g \in \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$ amounts to assigning the finitely many continuous functions $f_{g,m} \in C(\mathbb{H})$ with $m \in M_2(\mathbb{Z}/N\mathbb{Z})$, where we write

$$f(g,\rho,z) = f_{g,p_N(\rho)}(z).$$

The invariance condition

$$f(\gamma_1 g\gamma_2,\rho,z) = f(g,\gamma_2\rho,\gamma_2(z)), \tag{3.328}$$

for all $\gamma_j \in \Gamma$, $\forall g \in \mathrm{GL}_2^+(\mathbb{Q})$ and for all $(\rho,z) \in M_2(\hat{\mathbb{Z}})\times\mathbb{H}$, then shows (cf. Lemma 3.42) that

$$f_{g,m}|\gamma = f_{g,m}, \quad \forall\gamma \in \Gamma(N)\cap g^{-1}\Gamma g, \tag{3.329}$$

with standard notations for congruence subgroups and for the slash operation in weight 0.

As above, we denote by F the modular field, i.e. the union of the fields F_N of modular functions of level N rational over $\mathbb{Q}(e^{2\pi i/N})$.

LEMMA 3.81. *Let $f \in C(\mathcal{U}_2)$ have level N and assume that $f_{g,m} \in F$ for all (g,m). For any $\alpha \in \mathrm{SL}_2(\hat{\mathbb{Z}})$ one has*

$$f_{g,\alpha m} = \mathrm{Gal}(\alpha) f_{g',m} \tag{3.330}$$

where $g\alpha = \alpha' g'$ is the decomposition of $g\alpha$ as an element in the product $\mathrm{GL}_2(\hat{\mathbb{Z}})\mathrm{GL}_2(\mathbb{Q})$.

PROOF. Notice that the decomposition $\alpha' g'$ is not unique, and one can replace the pair (α', g') by $(\alpha'\gamma,\ \gamma^{-1}g')$ for any $\gamma \in \Gamma$. But the left invariance (3.328) implies

$$f(\gamma^{-1}\, g', \rho, \tau) = f(g', \rho, \tau)$$

for all $\gamma \in \Gamma$ so that the condition (3.330) is well defined. Consider the map

$$(g, \alpha) \in \mathrm{GL}_2^+(\mathbb{Q}) \times \mathrm{GL}_2(\hat{\mathbb{Z}}) \mapsto g'(g, \alpha) \in \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q}) \tag{3.331}$$

where g' is the unique class with $g\alpha = \alpha' g'$ for some $\alpha' \in \mathrm{GL}_2(\hat{\mathbb{Z}})$. One has

$$g'(g, \alpha) = p(g\alpha)$$

where p is the projection of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ onto the quotient

$$\mathrm{GL}_2(\hat{\mathbb{Z}})\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \sim \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q}).$$

Since p is continuous for the discrete topology of $\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$, so is the map g'. It follows in particular that for fixed $g \in \mathrm{GL}_2^+(\mathbb{Q})$ there are only finitely many possible values of $g'(g,\alpha)$ when α varies in the compact space $\mathrm{GL}_2(\hat{\mathbb{Z}})$. This shows that both sides of (3.330) depend continuously on α and by the density of Γ in $\mathrm{SL}_2(\hat{\mathbb{Z}})$ it is enough to prove (3.330) for $\alpha = \gamma \in \Gamma$. Then we have $g' = g\gamma$ and

$$f_{g,\alpha m}(\tau) = f(g, \gamma m, \tau) = f(g\gamma, m, \gamma^{-1}(\tau))$$

and

$$f_{g,\alpha m} = f_{g',m}|\gamma^{-1}$$

Thus one obtains the required condition using Theorem 3.75 a). □

We now show that the basic examples $\mathcal{X}_a$, $\mathcal{E}_{2m,a}$ and T_{r_1,r_2} of arithmetic elements fulfill a condition finer than (3.330) in that $\mathrm{SL}_2(\hat{\mathbb{Z}})$ is replaced by $\mathrm{GL}_2(\hat{\mathbb{Z}})$.

PROPOSITION 3.82. *The $\mathcal{X}_a$, $\mathcal{E}_{2k,a}$ and the T_{r_1,r_2} fulfill the following "arithmetic" conditions*

- *f has finite support in the variable $g \in \Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})$.*
- *f has finite level in $\rho \in M_2(\hat{\mathbb{Z}})$*
- *The functions $f_{g,m}(z)$, for $m \in M_2(\mathbb{Z}/N\mathbb{Z})$, fulfill*

$$f_{g,m} \in F \quad \forall (g,m). \tag{3.332}$$

- *Let $\alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ and $g\,\alpha = \alpha'\, g'$ be the decomposition of $g\,\alpha$ as a product in $\mathrm{GL}_2(\hat{\mathbb{Z}})\mathrm{GL}_2(\mathbb{Q})$, then*

$$f_{g,\alpha m} = \mathrm{Gal}(\alpha) f_{g',m} \qquad \forall \alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})\,. \tag{3.333}$$

PROOF. Let us first consider the $\mathcal{X}_a$ (the case of the $\mathcal{E}_{2k,a}$ is similar). The support in the variable $g \in \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})$ of the function $f = \mathcal{X}_a$ defined by (3.327) has one element. For $a \in (\frac{1}{N}\mathbb{Z}/\mathbb{Z})^2$, $f = \mathcal{X}_a$ has level N since it only depends upon $\rho(a)$ i.e. on the restriction $p_N(\rho)$ of ρ to N-torsion. The corresponding functions $f_{g,m}$ are of the form $\mathfrak{f}_u$ by construction of the $\mathcal{X}_a$. In fact, one has

$$f_{1,m} = \mathfrak{f}_{m(a)} \quad \forall m. \tag{3.334}$$

Thus, Theorem 3.75 b) allows one to check the last condition for the $\mathcal{X}_a$. Notice that (3.334) only holds if $m(a) \neq 0$ and we still have to handle the case when $m(a) = 0$. In that case one has $\alpha\, m(a) = 0$ and both sides vanish by construction of $\mathcal{X}_a$. In the case of the $\mathcal{E}_{2k,a}$ both sides are equal to ν_{2k} but the latter is invariant under the Galois action of $\mathrm{GL}_2(\hat{\mathbb{Z}})$ on F so that one gets the required equality.
Let us now deal with the T_{r_1,r_2}. One just needs to check that the functions (3.326) fulfill (3.333). This follows since the map (3.331) fulfills:

$$g'(g,\alpha) = g \quad \text{in } \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})/\Gamma\,, \quad \forall g \in \mathrm{GL}_2^+(\mathbb{Q})\,. \tag{3.335}$$

To prove (3.335) one can assume that g is diagonal since every double class contains a diagonal element. One can further factorize α as

$$\alpha = \delta\alpha_0\,, \quad \delta = \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix}, \quad \alpha_0 \in \mathrm{SL}_2(\hat{\mathbb{Z}})\,.$$

This allows one to assume that $\alpha \in \mathrm{SL}_2(\hat{\mathbb{Z}})$, in which case (3.335) follows by the density of $\Gamma \subset \mathrm{SL}_2(\hat{\mathbb{Z}})$. $\square$

We can now define "arithmetic" requirements that characterize $f \in \mathcal{A}_{2,\mathbb{Q}}$. Namely we require the following.

DEFINITION 3.83. *A function $f \in C(\mathcal{U}_2)$ is* arithmetic *if and only if it satisfies the conditions of Proposition 3.82. We let $\mathcal{A}_{2,\mathbb{Q}}$ be the linear space of arithmetic elements $f \in C(\mathcal{U}_2)$.*

REMARK 3.84. The condition (3.333) of Proposition 3.82 is automatic on $\mathrm{SL}_2(\hat{\mathbb{Z}}) \subset \mathrm{GL}_2(\hat{\mathbb{Z}})$ by Lemma 3.81. It is however crucial to impose the stronger condition in this form. In particular (3.333) shows that functions on $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ extended as above by 0 outside $\Gamma \times M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ and which are independent of the variable $\rho \in M_2(\hat{\mathbb{Z}})$ are *arithmetic* iff they belong to F and are invariant under $\mathrm{GL}_2(\hat{\mathbb{Z}})$ and hence iff they belong to $\mathbb{Q}[j]$. Elements of F fixed under $\mathrm{SL}_2(\hat{\mathbb{Z}})$ coincide with $\mathbb{Q}^{\mathrm{ab}}[j]$, where $\mathbb{Q}^{\mathrm{ab}}$ is the cyclotomic extension of $\mathbb{Q}$.
In fact, as we have seen in §6 above (cf. [**268**]), the modular field F_N of level N contains a primitive N-th root of 1. Thus, the condition (3.332) alone would allow the algebra $\mathcal{A}_{2,\mathbb{Q}}$ to contain the cyclotomic field $\mathbb{Q}^{\mathrm{ab}} \subset \mathbb{C}$. However, this would prevent the existence of any intertwining between a Galois action on values of states and the action of symmetries on the algebra, which is one of the main goals of this quantum statistical mechanical formulation.
In fact, the existence of an intertwining

$$\theta : S = \mathbb{Q}^*\backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \xrightarrow{\cong} \mathrm{Gal}(F_\varphi/\mathbb{Q})$$

with

$$\alpha \circ \varphi(f) = \varphi \circ \theta^{-1}(\alpha)(f), \quad \forall \alpha \in \mathrm{Gal}(F_\varphi/\mathbb{Q}), \quad \forall f \in \mathcal{A}_{2,\mathbb{Q}} \tag{3.336}$$

implies that any $f \in \mathcal{A}_{2,\mathbb{Q}} \cap \mathbb{C}$ is fixed under the Galois action and hence belongs to $\mathbb{Q}$. This follows immediately from the $\mathbb{C}$-linearity of the functional φ. The condition (3.333) on the elements of the arithmetic algebra $\mathcal{A}_{2,\mathbb{Q}}$ has the effect that, if the spectrum of an element of $\mathcal{A}_{2,\mathbb{Q}}$ contains a root of unity, then it also has to contain all of its Galois conjugates, so that, in particular, no such element can be a scalar $\notin \mathbb{Q}$.

We now show that the arithmeticity condition (3.333) of Definition 3.83 can be formulated as a consistency condition on the roots of unity that appear in the coefficients of the q-series, when $\rho \in M_2(\hat{\mathbb{Z}})$ is multiplied on the left by a diagonal matrix.

Consider elements $g \in \mathrm{GL}_2^+(\mathbb{Q})$ and $\alpha \in \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, respectively of the form

$$g = r\begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.337}$$

with k prime to N and $n|N$.

Lemma 3.85. *An element $f \in C(\mathcal{U}_2)$ is arithmetic if and only if it satisfies the first two conditions of Proposition 3.82 and the following property. If f is of level N_0, then for any multiple N of N_0 and any pair (g, α) as in* (3.337) *we have $f_{g,m} \in F_N$ for all $m \in M_2(\mathbb{Z}/N\mathbb{Z})$ and the q-series of $f_{g,\alpha m}$ is obtained from the q-series for $f_{g,m}$ by raising to the power k the roots of unity that appear as coefficients.*

Proof. Let f fulfill the condition of the Lemma. One has $f_{g,m} \in F_N \subset F$ so we only need to check the covariance condition (3.333). By Lemma 3.81, the only nontrivial part of (3.333) is the case of diagonal matrices

$$\delta = \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix}$$

with $u \in \mathrm{GL}_1(\hat{\mathbb{Z}})$. To prove (3.333) we can assume that $g = g_0\gamma$, with g_0 diagonal as in (3.337) and $\gamma \in \Gamma$ since every double class contains a diagonal element. Let then $\gamma\alpha = \delta\alpha_1$ with δ as above and $\alpha_1 \in \mathrm{SL}_2(\hat{\mathbb{Z}})$. One has, using the condition of the Lemma,

$$f_{g_0,\delta\alpha_1 m} = \mathrm{Gal}(\delta) f_{g_0,\alpha_1 m}$$

since, by Lemma 3.76, $\mathrm{Gal}(\delta)$ acts by raising the roots of unity that appear as coefficients of the q-expansion to the power k, where u is the residue of k modulo N (cf. [**268**] (6.2.1) p.141). One then has, using $g = g_0\gamma$ and $\gamma\alpha = \delta\alpha_1$,

$$f_{g,\alpha m} = f_{g_0,\gamma\alpha m}|\gamma = \mathrm{Gal}(\gamma^{-1}) f_{g_0,\gamma\alpha m} = \mathrm{Gal}(\gamma^{-1}\delta)\, f_{g_0,\alpha_1 m}$$

and by Lemma 3.81, with $g_0\alpha_1 = \alpha_1' g_0'$ we get

$$f_{g,\alpha m} = \mathrm{Gal}(\gamma^{-1}\delta\alpha_1)\, f_{g_0',m} = \mathrm{Gal}(\alpha)\, f_{g_0',m}.$$

Moreover, one also has

$$g\alpha = g_0\gamma\alpha = g_0\delta\alpha_1 = \delta g_0\alpha_1 = \delta\alpha_1' g_0',$$

which shows that $g' = g_0'$. One then checks the converse similarly. $\square$

We now show that $\mathcal{A}_{2,\mathbb{Q}}$ is a subalgebra of the algebra of unbounded multipliers and there is an induced action on the algebra $\mathcal{A}_{2,\mathbb{Q}}$ of the symmetries of $(\mathcal{A}_2, \sigma_t)$ analyzed in §5.4.

PROPOSITION 3.86. *The space $\mathcal{A}_{2,\mathbb{Q}}$ is a subalgebra of the algebra of unbounded multipliers of the C^*-algebra $\mathcal{A}_2$, which is globally invariant under the action of the symmetry group $S = \mathbb{Q}^*\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ of $(\mathcal{A}_2, \sigma_t)$.*

PROOF. We say that a point $\tau \in \mathbb{H}$ is generic if $j(\tau)$ is transcendental. We have seen that, for each generic $\tau \in \mathbb{H}$, the evaluation map $h \mapsto I_\tau(h) \in \mathbb{C}$ which is given by $I_\tau(h) = h(\tau)$ on regular elements gives an isomorphism of F with a subfield $F_\tau \subset \mathbb{C}$ and a corresponding action Gal_τ of $\mathrm{GL}_2(\mathbb{A}_f)$ by automorphisms of F_τ, such that

$$\mathrm{Gal}_\tau(g)(I_\tau(h)) = I_\tau(\mathrm{Gal}(g)(h)). \tag{3.338}$$

The convolution product in $\mathcal{A}_2$ and on multipliers is given by

$$(f_1 * f_2)(g,\rho,\tau) = \sum_{g_1\in\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q}):\, g_1\rho\in M_2(\hat{\mathbb{Z}})} f_1(gg_1^{-1}, g_1\rho, g_1(\tau)) f_2(g_1,\rho,\tau)\,.$$

which means, in terms of the $f_{j,g}$ as in (3.200),

$$(f_1 * f_2)_g = \sum_{s\in\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q})} f_{1,gs^{-1}}|s\ f_{2,s}.$$

It follows using Lemma 3.74 that if f_1 and f_2 are arithmetic elements then $(f_1 * f_2)_{g,m} \in F$ (compare with Proposition 2 of [**93**]). It remains to show that condition (3.333) is stable under convolution, i.e. that for any two elements f_1 and f_2 in $\mathcal{A}_{2,\mathbb{Q}}$ the convolution product $f_1 * f_2$ still fulfills (3.333). Recall that given $\alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ and $g \in \mathrm{GL}_2(\mathbb{Q})$ there are elements $g' \in \mathrm{GL}_2(\mathbb{Q})$ and $\beta \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ such that $g\alpha = \beta g'$. By definition, one has

$$(f_1 * f_2)(g,\alpha\rho,\tau) = \sum_{g_1\in\Gamma\backslash\mathrm{GL}_2^+(\mathbb{Q}):\, g_1\alpha\rho\in M_2(\hat{\mathbb{Z}})} f_1(gg_1^{-1}, g_1\alpha\rho, g_1(\tau)) f_2(g_1,\alpha\rho,\tau).$$

We let $g_1\alpha = \alpha' g_1'$ be the decomposition of $g_1\alpha$ as above. We then use (3.333) to write the right hand side as

$$(f_1 * f_2)(g,\alpha\rho,\tau) = \sum f_1(gg_1^{-1}, g_1\alpha\rho, g_1(\tau))\mathrm{Gal}_\tau(\alpha)(f_2(g_1',\rho,\tau)),$$

with Gal_τ as in (3.338). The result then follows from the equality

$$f_1(gg_1^{-1}, g_1\alpha\rho, g_1(\tau)) = \mathrm{Gal}_\tau(\alpha)(f_1(g'{g_1'}^{-1}, g_1'\rho, g_1'(\tau))), \tag{3.339}$$

which we now prove. The equality $g_1\alpha = \alpha' g_1'$ together with (3.333) shows that

$$f_1(gg_1^{-1}, g_1\alpha\rho, g_1(\tau)) = \mathrm{Gal}_{g_1(\tau)}(\alpha')(f_1(g'{g_1'}^{-1}, g_1'\rho, g_1(\tau))),$$

using $gg_1^{-1}\alpha' = gg_1^{-1}(g_1\alpha{g_1'}^{-1}) = g\alpha{g_1'}^{-1} = \beta g'{g_1'}^{-1}$.
For any $h \in F$, one has

$$I_{g_1(\tau)}(\mathrm{Gal}(\alpha')h) = I_{g_1(\tau)}(\mathrm{Gal}(g_1)\mathrm{Gal}(\alpha)\mathrm{Gal}({g_1'}^{-1})(h))$$

and, by construction of the Galois action (cf. Theorem 3.75 a)), we have

$$I_{g_1(\tau)} \circ \mathrm{Gal}(g_1) = I_\tau\,, \quad \forall g_1 \in \mathrm{GL}_2^+(\mathbb{Q})\,, \tag{3.340}$$

so that in fact we obtain

$$\mathrm{Gal}_{g_1(\tau)}(\alpha')I_{g_1(\tau)}(h) = \mathrm{Gal}_\tau(\alpha)I_{g_1'(\tau)}(h).$$

This proves (3.339) and it shows that $\mathcal{A}_{2,\mathbb{Q}}$ is a subalgebra of the algebra of unbounded multipliers of $\mathcal{A}_2$. To prove the invariance under $S = \mathbb{Q}^*\backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ is straightforward, since the endomorphisms are all acting on the ρ variable by right multiplication, which does not interfere with condition (3.333). ☐

PROPOSITION 3.87. *The map*

$$\xi(f)(g,\rho,\tau) := \begin{cases} f(g) & g \in M_2^+(\mathbb{Z}) \\ 0 & \textit{otherwise} \end{cases}$$

defines an isomorphism

$$\xi : \mathcal{H}(M_2^+(\mathbb{Z}),\Gamma) \to \mathcal{A}_{2,\mathbb{Q}} \tag{3.341}$$

of the standard Hecke algebra $\mathcal{H}(M_2^+(\mathbb{Z}),\Gamma)$ of Γ-bi-invariant functions with Γ-finite support on $M_2^+(\mathbb{Z})$, with a subalgebra $\mathcal{H} \subset \mathcal{A}_{2,\mathbb{Q}}$.

PROOF. The convolution product in the Hecke algebra is given on Γ-bi-invariant functions with Γ-finite support by

$$(f_1 * f_2)(g) = \sum_{g_1\in\Gamma\backslash M_2^+(\mathbb{Z}):\, gg_1^{-1}\in M_2^+(\mathbb{Z})} f_1(gg_1^{-1})f_2(g_1) \tag{3.342}$$

One has $(\xi(f_1) * \xi(f_2))(g,\rho,\tau)$ given by the expression

$$\sum_{g_1\in\Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q}):\, g_1\rho\in M_2(\hat{\mathbb{Z}})} \xi(f_1)(gg_1^{-1}, g_1\rho, g_1(\tau))\xi(f_2)(g_1,\rho,\tau).$$

The term in the sum vanishes unless $g_1 \in M_2^+(\mathbb{Z})$ and $gg_1^{-1} \in M_2^+(\mathbb{Z})$. This implies $g \in M_2^+(\mathbb{Z})$ so that the support of $\xi(f_1)*\xi(f_2)$ is contained in $M_2^+(\mathbb{Z})$. One can thus assume that $g \in M_2^+(\mathbb{Z})$ and restrict the sum to the subset of $g_1 \in \Gamma\backslash M_2^+(\mathbb{Z})$ with $gg_1^{-1} \in M_2^+(\mathbb{Z})$. Notice that the condition $g_1\rho \in M_2(\hat{\mathbb{Z}})$ in the convolution product is now automatically satisfied, so that ξ is a homomorphism. ☐

7.3. Division relations in the arithmetic algebra.

We now describe the algebraic relations fulfilled by the $\mathcal{X}_a$ as extensions of the division formulas of elliptic functions. These extend the division relations of elliptic functions from invertible $\mathbb{Q}$-lattices to arbitrary ones.
In order to obtain the division formulas for the Eisenstein series $E_{2m,a}$ we need to control the image of $\left(\frac{1}{N}\mathbb{Z}\right)^2 = \frac{1}{N}\mathbb{Z}^2$ under an arbitrary element $\rho \in M_2(\hat{\mathbb{Z}})$. This is done as follows using projections q_Δ indexed by sublattices $\Delta \subset \mathbb{Z}^2$ associated to the closed and open subsets $m\,M_2(\hat{\mathbb{Z}})$ for $m \in M_2^+(\mathbb{Z})$. As in Lemma 3.53 the subset $L_m = m\,M_2(\hat{\mathbb{Z}})$ only depends upon the subgroup $m(\mathbb{Z}^2) \subset \mathbb{Z}^2$ and we let

$$q_\Delta = \mathbf{1}_{L_m}\,, \quad \forall m \in M_2^+(\mathbb{Z})\,, \quad m(\mathbb{Z}^2) = \Delta\,. \tag{3.343}$$

One has, for any $m \in M_2^+(\mathbb{Z})$ with $\Delta = m\mathbb{Z}^2$,

$$q_\Delta(\rho) = 1 \Leftrightarrow\; m^{-1}\rho \in M_2(\hat{\mathbb{Z}})\,. \tag{3.344}$$

Note that the q_Δ are not Γ-invariant on the left. In fact one has

$$q_\Delta(\gamma\rho) = q_{\gamma^{-1}\Delta}(\rho)\,, \quad \forall\gamma \in \Gamma\,,\; \rho \in M_2(\hat{\mathbb{Z}})\,. \tag{3.345}$$

One has

$$\Delta \subset \Delta' \Rightarrow q_\Delta \leq q_{\Delta'} \,. \tag{3.346}$$

Indeed $m\mathbb{Z}^2 \subset m'\mathbb{Z}^2$ shows that $(m')^{-1}m\mathbb{Z}^2 \subset \mathbb{Z}^2$ and $m = m'm''$ for some $m'' \in M_2^+(\mathbb{Z})$ so that $L_m \subset L_{m'}$.
The projections q_Δ commute and for any pair of lattices $\Delta, \Delta' \subset \mathbb{Z}^2$ one has

$$q_\Delta \, q_{\Delta'} = q_{\Delta\cap\Delta'} \,. \tag{3.347}$$

This follows from the second part of the next lemma:

LEMMA 3.88. *Let $N \in \mathbb{N}$, and $\rho \in M_2(\hat{\mathbb{Z}})$.*

(1) *There exists $\gamma \in \Gamma$, $d_j \mid N$ with $d_1 \mid d_2$ and $\mu \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ such that*

$$\rho = \gamma \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \mu \quad \textit{modulo } N\, M_2(\hat{\mathbb{Z}}) \tag{3.348}$$

(2) *Let $\Delta \subset \mathbb{Z}^2$ with $\Delta \supset N\mathbb{Z}^2$. One has*

$$q_\Delta(\rho) = 1 \Leftrightarrow \rho(\frac{1}{N}\mathbb{Z}^2) \subset \frac{1}{N}\Delta \quad \textit{modulo } \mathbb{Z}^2 \,. \tag{3.349}$$

PROOF. (1) The statement only depends on the projection $p_N(\rho) = \rho_N \in M_2(\mathbb{Z}/N\mathbb{Z})$. Thus one can lift ρ to an element of $M_2^+(\mathbb{Z})$ and write the latter as a product

$$\gamma_1 \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \gamma_2$$

where $\gamma_j \in \Gamma$ and $a_j \in \mathbb{N}$, $a_j > 0$ with $a_1 \mid a_2$. One has $a_j\,\mathbb{Z}/N\mathbb{Z} = d_j\,\mathbb{Z}/N\mathbb{Z}$ for some unique divisors $d_j \mid N$, $d_j \in \{1, \ldots, N\}$ and $d_1 \mid d_2$ by construction. Thus one can lift a_j modulo N in the form $d_j u_j$, where $u_j \in \mathrm{GL}_1(\hat{\mathbb{Z}})$. This gives the required lift of ρ of the form (3.348).
(2) Let Δ with $N\mathbb{Z}^2 \subset \Delta \subset \mathbb{Z}^2$ be such that $q_\Delta(\rho) = 1$. Let us show that $\rho(\frac{1}{N}\mathbb{Z}^2) \subset \frac{1}{N}\Delta$. Let $m \in M_2^+(\mathbb{Z})$ be such that $\Delta = m\mathbb{Z}^2$. Then $q_\Delta(\rho) = 1$ implies that $\rho = m\mu$ for some $\mu \in M_2(\hat{\mathbb{Z}})$. Thus $\rho(\frac{1}{N}\mathbb{Z}^2) \subset m(\frac{1}{N}\mathbb{Z}^2) = \frac{1}{N}\Delta$. Conversely let Δ', $N\mathbb{Z}^2 \subset \Delta' \subset \mathbb{Z}^2$ be defined by $\rho(\frac{1}{N}\mathbb{Z}^2) = \frac{1}{N}\Delta'$ modulo $\mathbb{Z}^2$. We need to show that $q_{\Delta'}(\rho) = 1$, i.e. that there exists $m' \in M_2^+(\mathbb{Z})$ such that $\Delta' = m'\mathbb{Z}^2$ and ${m'}^{-1}\rho \in M_2(\hat{\mathbb{Z}})$. By the first part of the lemma, one can assume that ρ is of the form (3.348). One then takes

$$m' = \gamma \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \in M_2^+(\mathbb{Z})$$

and checks that $\Delta' = m'\mathbb{Z}^2$ while ${m'}^{-1}\rho$ belongs to $M_2(\hat{\mathbb{Z}})$. $\square$

Given an integer $N > 1$, we let S_N be the set of lattices

$$N\mathbb{Z}^2 \subset \Delta \subset \mathbb{Z}^2. \tag{3.350}$$

This is the same as the set of subgroups of $(\mathbb{Z}/N\mathbb{Z})^2$. For each $\Delta \in S_N$ we define a projection $\pi(N, \Delta)$ by

$$\pi(N, \Delta) = q_\Delta \prod_{\Delta' \in S_N :\, \Delta' \subsetneq \Delta} (1 - q_{\Delta'}) \,. \tag{3.351}$$

By Lemma 3.88 the range of $\pi(N,\Delta)$ is exactly the set of $\rho \in M_2(\hat{\mathbb{Z}})$ such that

$$\rho(\frac{1}{N}\mathbb{Z}^2) = \frac{1}{N}\Delta \ \text{ modulo } \mathbb{Z}^2 . \tag{3.352}$$

The general form of the division relations is then as follows.

PROPOSITION 3.89. *There exist canonical modular forms $\omega_{N,\Delta,k}$ of level N and weight $2k$, such that for all k and $(\rho,\tau) \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ they satisfy*

$$\sum_{Na=0} X_a^k(\rho,\tau) = \sum_{\Delta \in S_N} \pi(N,\Delta)(\rho)\, \omega_{N,\Delta,k}(\tau) , \tag{3.353}$$

and

$$\omega_{N,\Delta,k}(\gamma\tau) = (c\tau+d)^{2k}\, \omega_{N,\gamma^{-1}\Delta,k}(\tau) , \quad \forall \gamma \in \Gamma , \tag{3.354}$$

which implies that $\omega_{N,\Delta,k}$ is of level N.

PROOF. We prove the result first for $k=1$ and then we proceed by induction on k.

The division formulas in weight 2 involve the 1-cocycle on the group $\mathrm{GL}_2^+(\mathbb{Q})$ with values in Eisenstein series of weight 2, given in terms of the Dedekind η-function by (cf. [**93**])

$$\mu_\gamma(\tau) = \frac{1}{2\pi i}\frac{d}{d\tau}\log\frac{\eta^4|\gamma}{\eta^4}, \tag{3.355}$$

where we used the standard 'slash operator' notation (3.248). Since $\mu_\gamma = 0$ for $\gamma \in \Gamma$, the cocycle property

$$\mu_{\gamma_1\gamma_2} = \mu_{\gamma_1}|\gamma_2 + \mu_{\gamma_2} \tag{3.356}$$

shows that, for $m \in M_2^+(\mathbb{Z})$, the value of $\mu_{m^{-1}}$ only depends upon the lattice $\Delta = m(\mathbb{Z}^2)$. We shall denote it by μ_Δ.

We let, with $q = e^{2\pi i\tau}$, $\Im\tau > 0$,

$$e_2(\tau) = -\frac{1}{12} + 2\sum_{m,n>0} n\, q^{nm} . \tag{3.357}$$

With the notations of [**93**] one has

$$e_2 = (2\pi i)^{-2}\, G_2^* ,$$

so that by [**93**] (3.11) one gets, for any $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$,

$$\mu_\gamma(\tau) = -2\,(e_2|\gamma(\tau) - e_2(\tau)) - \frac{2i}{\pi}\frac{c}{c\tau+d} , \quad \forall\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} . \tag{3.358}$$

The case $k=1$ is taken care of by the following lemma.

LEMMA 3.90. *For any integer N, the X_a with $a \in \mathbb{Q}/\mathbb{Z}$ fulfill the relation*

$$\sum_{N\,a=0} X_a = -\frac{1}{2}N^2 \sum_{\Delta\in S_N} \pi(N,\Delta)\,\mu_\Delta .$$

PROOF. By construction, the projections $\pi(N,\Delta)$ with $\Delta \in S_N$ form a partition of unity, namely

$$\sum_{\Delta \in S_N} \pi(N,\Delta) = 1.$$

Thus, to prove the lemma it is enough to evaluate both sides on a ρ in the support of $\pi(N,\Delta)$.
We can then use the equality

$$\mu_{\gamma^{-1}\Delta} = \mu_\Delta|\gamma, \quad \forall \gamma \in \Gamma \tag{3.359}$$

to assume, using Lemma 3.88, that Δ and ρ_N are of the form

$$\Delta = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \mathbb{Z}^2, \quad \text{and} \quad \rho_N = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}, \tag{3.360}$$

where $d_j \mid N$ and $d_1 \mid d_2$. Let $d_2 = nd_1$. The order of the kernel of ρ_N is $d_1 d_2$. Recall that $X_a(\rho,\tau)$ is defined by (3.318) where the $\mathbb{Q}$-lattice $L(\rho,\tau)$ is given by (3.314), i.e.

$$L(\rho,\tau) = (\Lambda,\phi), \quad \Lambda = \mathbb{Z} + \tau\mathbb{Z} \quad \text{and} \quad \phi(a) = \rho_1(a) - \tau\rho_2(a) \in \mathbb{Q}\Lambda/\Lambda.$$

Thus with the above notations we get,

$$\phi(a) = \frac{d_1 a_1}{N} - \frac{d_2 a_2}{N}\tau, \quad \forall a = (\frac{a_1}{N}, \frac{a_2}{N}) \in \frac{1}{N}\mathbb{Z}^2.$$

For $a \in (\frac{1}{N}\mathbb{Z}^2/\mathbb{Z}^2)/\mathrm{Ker}\,\rho_N$ the cosets $\phi(a) + \Lambda$ are disjoint and

$$\Lambda' = \cup_a(\phi(a) + \Lambda) = \frac{d_1}{N}(\mathbb{Z} + n\tau\mathbb{Z})$$

Thus the computation of $\sum_{Na=0} X_a(\rho,\tau)$ gives

$$(2\pi i)^{-2}\, d_1\, d_2 \sum_{\Lambda' \smallsetminus \{0\}} y^{-2} - N^2 e_2(\tau),$$

using (3.318) even for $\phi(a) \in \Lambda$ just by removing the term y^{-2} for $y = 0$ from the sum. This gives $N^2(n\, e_2(n\,\tau) - e_2(\tau)) = -\frac{1}{2}N^2\mu_\Delta$ using (3.358). ☐

This proves, using (3.359), the proposition for $k = 1$. For the general case, we can proceed by induction using (3.320) to express X_a^k as $E_{2k,a}$ plus a polynomial of degree $< k$ in X_a with coefficients in $\mathbb{Q}[e_4, e_6]$. Thus, we only need to prove the equality

$$\sum_{Na=0} E_{2k,a} = \sum_{\Delta \in S_N} \pi(N,\Delta)\alpha_{N,\Delta,k}, \tag{3.361}$$

where the modular forms $\alpha_{N,\Delta,k}$ are given explicitly as

$$\alpha_{N,\Delta,k} = \det(m)\,(N^{2k} n^k e_{2k}|m^{-1} - e_{2k} + \nu_{2k}), \tag{3.362}$$

with $m \in M_2^+(\mathbb{Z})$ such that $\Delta = m(\mathbb{Z}^2)$, and $(d_1, d_2) = (d, dn)$ associated to Δ as above (cf. Lemma 3.88). One checks that

$$\alpha_{N,\Delta,k}|\gamma = \alpha_{N,\gamma^{-1}\Delta,k}, \quad \forall \gamma \in \Gamma. \tag{3.363}$$

Thus one can proceed as in the proof of Lemma 3.90 and obtain (3.361), by evaluating both sides on arbitrary $\rho \in \pi(N, \Delta)$ with ρ and Δ as in (3.360). One has

$$\sum_{Na=0} E_{2k,a}(\rho, \tau) = d_1\, d_2\, ((2\pi i)^{-2k} \sum_{\Lambda' \smallsetminus \{0\}} y^{-2k} - e_{2k}(\tau) + \nu_{2k}(\tau))$$

$$= \det(m)\, (\frac{N^{2k}}{d^{2k}} e_{2k}(n\tau) - e_{2k}(\tau) + \nu_{2k}(\tau))$$

where we took care of replacing the contribution of the coset $\Lambda \subset \Lambda'$ corresponding to $\phi(a) \in \Lambda$ by the value dictated by (3.319). One has $\det(m^{-1})^k = d^{-2k} n^{-k}$. Thus one gets (3.362). $\square$

One can rewrite the relation (3.353) in terms of the weight 0 elements $\mathcal{X}_a = \mathbf{c}\, X_a$ as

$$\sum_{Na=0} \mathcal{X}_a^k(\rho, \tau) = \sum_{\Delta \in S_N} \pi(N, \Delta)(\rho)\, \beta_{N,\Delta,k}(\tau)\,, \tag{3.364}$$

where the $\beta_{N,\Delta,k} = \mathbf{c}^k\, \omega_{N,\Delta,k}$ fulfill

$$\beta_{N,\Delta,k}(\gamma\tau) = \beta_{N,\gamma^{-1}\Delta,k}(\tau)\,, \quad \forall \gamma \in \Gamma\,. \tag{3.365}$$

Note that the terms $\pi(N, \Delta)\, \beta_{N,\Delta,k}$ are not invariant under the left action of Γ on $M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ and hence do not make sense as elements of $\mathcal{A}_{2,\mathbb{Q}}$, but the sum

$$\sigma_k = \sum_{\Delta \in S_N} \pi(N, \Delta)\, \beta_{N,\Delta,k} \tag{3.366}$$

is Γ-invariant since one has, using (3.345),

$$\pi(N, \Delta)(\gamma\rho) = \pi(N, \gamma^{-1}\Delta)(\rho)\,. \tag{3.367}$$

By construction one has $\beta_{N,\Delta,k} \in \mathcal{O}_F$.

THEOREM 3.91. *Let $N \in \mathbb{N}$. Then the subalgebra of $\mathcal{A}_{2,\mathbb{Q}}$ generated by the $\mathcal{X}_a$ for $Na = 0$ is reduced and finite dimensional over $\mathbb{Q}[j]$.*

PROOF. Let us start with the algebra generated by the σ_k of (3.366), for $1 \leq k \leq N^2$. To show that it is finite-dimensional over $\mathbb{Q}[j]$ it is enough to show that each of the generators σ_k is integral over $\mathbb{Q}[j]$. By Proposition 3.72 the $\beta_{N,\Delta,k}$ are integral over $\mathbb{Q}[j]$. Thus there exists a monic polynomial $P(X)$ with coefficients in $\mathbb{Q}[j]$ such that

$$P(\sum \epsilon_\Delta\, \beta_{N,\Delta,k}) = 0$$

for any $\epsilon_\Delta \in \{0, 1\}$. It follows then from (3.366) that $P(\sigma_k) = 0$. By (3.364) the symmetric functions of the $\mathcal{X}_a$ for $Na = 0$ are polynomials in the σ_k and are thus integral over $\mathbb{Q}[j]$. We thus get that the generators $\mathcal{X}_a$ are integral over $\mathbb{Q}[j]$ and it follows that the subalgebra of $\mathcal{A}_{2,\mathbb{Q}}$ generated by the $\mathcal{X}_a$ for $Na = 0$ is finite-dimensional over $\mathbb{Q}[j]$. It does not contain nilpotent elements since it is by construction a subalgebra of the algebra of complex-valued continuous functions on $M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$. $\square$

7.4. KMS states.

In this section we describe the structure and arithmetic properties of KMS states of the system $(\mathcal{A}_2, \sigma_t)$ introduced in §5. We first review some terminology.

DEFINITION 3.92. *A positive energy representation of a quantum statistical mechanical system $(\mathcal{A}, \sigma_t)$ is a covariant representation (Definition 3.1)*

$$\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$$

of the C^*-algebra $\mathcal{A}$ as bounded operators on a Hilbert space $\mathcal{H}$ such that the time evolution σ_t is implemented:

$$\sigma_t(\pi(a)) = e^{itH}\pi(a)e^{-itH},$$

by a positive Hamiltonian H with discrete spectrum.

Before proceeding to the classification of low-temperature KMS states of the system, we need some preliminary results. The following lemma shows that, in essence, invertible $\mathbb{Q}$-lattices are ground states of the system.

LEMMA 3.93. (1) *Let $s : \mathbb{Q}^2/\mathbb{Z}^2 \to \mathbb{Q}^2$ be a section of the projection $\pi : \mathbb{Q}^2 \to \mathbb{Q}^2/\mathbb{Z}^2$. Then the set of elements*

$$s(a+b) - s(a) - s(b), \qquad a, b \in \mathbb{Q}^2/\mathbb{Z}^2 \tag{3.368}$$

generate $\mathbb{Z}^2$.

(2) *Let $L = (\Lambda, \phi)$ be an invertible $\mathbb{Q}$-lattice. If $L' = (\Lambda', \phi')$ is a $\mathbb{Q}$-lattice commensurable to L, then $\Lambda \subset \Lambda'$.*

PROOF. (1) Let $\mathcal{I} \subset \mathbb{Z}^2$ be the subgroup generated by the (3.368). If $\mathcal{I} \neq \mathbb{Z}^2$, we can assume, possibly after a change of basis, that there is a prime number p such that $\mathcal{I} \subset p\mathbb{Z} \oplus \mathbb{Z}$. Thus, if we restrict s to the p-torsion in $\mathbb{Q}^2/\mathbb{Z}^2$ and multiply it by p we obtain a map

$$(a, b) \in \mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z} \mapsto \sigma(a, b) = p\, s(\frac{a}{p}, \frac{b}{p}) \in \mathbb{Z}^2$$

since $s(\frac{a}{p}, \frac{b}{p}) \in (\frac{a}{p}, \frac{b}{p}) + \mathbb{Z}^2$. This map gives a group homomorphism

$$\sigma \;:\; \mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p^2\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z},$$

which is a section of the projection

$$\mathbb{Z}/p^2\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z}$$

but this is impossible since in $\mathbb{Z}/p^2\mathbb{Z}$ any element of p-torsion maps to 0 under the projection $\mathbb{Z}/p^2\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$.

(2) Let s and s' denote, respectively, a lift of ϕ modulo Λ and of ϕ' modulo Λ'. Since by commensurability $\phi - \phi' = 0$ modulo $\Lambda + \Lambda'$ we have $s(a) - s'(a) \in \Lambda + \Lambda'$ for all $a \in \mathbb{Q}^2/\mathbb{Z}^2$. Thus, it is possible to correct s modulo Λ and s' modulo Λ' so that $s = s'$. Then, for all $a, b \in \mathbb{Q}^2/\mathbb{Z}^2$, one has $s(a+b) - s(a) - s(b) \in \Lambda \cap \Lambda'$. By the first statement of the lemma and the invertibility of ϕ we obtain that $\Lambda \cap \Lambda' = \Lambda$. □

We give a general characterization of KMS states for the system $(\mathcal{A}_2, \sigma_t)$, which shows that they are given by measures on the space of 2-dimensional $\mathbb{Q}$-lattices up to scaling. We then show that, for sufficiently low-temperature, these measures are supported on the commensurability classes of invertible $\mathbb{Q}$-lattices. This will be the key result for the classification of low-temperature KMS states.

We have the following preliminary result.

LEMMA 3.94. *Let* $X = \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ *and* $Y = M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$, *with* $p_\Gamma : Y \to X$ *the quotient map. Let* C *denote the closed subset of* X *given by* $C = p_\Gamma(\{(0, z)|z \in \mathbb{H}\})$. *Consider* $g \in \mathrm{GL}_2^+(\mathbb{Q})$, *such that* $g \notin \Gamma$ *and an element* $x \in X$, *with* $x \notin C$. *There exists a neighborhood* V *of* x *with the property that*

$$p_\Gamma(g\,p^{-1}(V)) \cap V = \emptyset \tag{3.369}$$

PROOF. The subset $\{(0, z)|z \in \mathbb{H}\} \subset M_2(\hat{\mathbb{Z}}) \times \mathbb{H}$ is closed and Γ-invariant, thus C is closed. The map $y \mapsto gy$ in Y does not define a map in X but a correspondence which assigns to a point $x \in X$ a finite subset of X. This correspondence is described in terms of a finite cover X_0 of X and two maps p_j from X_0 to X. Let $\Gamma_0 = \Gamma \cap g^{-1}\Gamma g$ and $X_0 = \Gamma_0\backslash Y$. For $x_0 \in X_0$, the projections $p_1(x_0) = p_\Gamma(y)$ and $p_2(x_0) = p_\Gamma(gy)$ are independent of the representative $y \in Y$ with $x_0 = \Gamma_0 y$. Moreover, if $p_1(x_0) \notin C$ then $p_1(x_0) \neq p_2(x_0)$ by Lemma 3.45. By construction Γ_0 is of finite index n in Γ and the fiber $p_1^{-1}(x)$ has at most n elements. Let then $W \subset X_0$ be a compact neighborhood of $p_1^{-1}(x)$ in X_0 such that $x \notin p_2(W)$. For $z \in V \subset X$ sufficiently close to x one has $p_1^{-1}(z) \subset W$, hence $p_2(p_1^{-1}(z)) \subset V^c$ which gives the result. $\square$

We can then obtain the following preliminary step in the characterization of KMS states of the GL_2 system $(\mathcal{A}_2, \sigma_t)$.

THEOREM 3.95. *For any given* $\beta > 0$, *let* φ *be a* KMS_β *state for the system* $(\mathcal{A}_2, \sigma_t)$. *Then there exists a probability measure* μ *on* $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ *with the property that*

$$\varphi(f) = \int_{\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})} f(1, \rho, z)\, d\mu(\rho, z), \qquad \forall f \in \mathcal{A}_2. \tag{3.370}$$

PROOF. Recall that elements $f \in \mathcal{A}_c(\mathcal{U}_2)$ are functions with compact support on the quotient of the space $\tilde{\mathcal{U}}_2$ of (3.195) by the action (3.193) of $\Gamma \times \Gamma$. Let $g_0 \in \mathrm{GL}_2^+(\mathbb{Q})$ and $f \in \mathcal{A}_c(\mathcal{U}_2)$ be such that

$$f(g, \rho, z) = 0, \quad \forall g \notin \Gamma g_0 \Gamma. \tag{3.371}$$

Any element $f \in \mathcal{A}_c(\mathcal{U}_2)$ is a finite linear combination of such functions.
It is then enough to show that we have $\varphi(f) = 0$ whenever $f \in \mathcal{A}_c(\mathcal{U}_2)$ satisfies the condition (3.371) with $g_0 \notin \Gamma$.
The formula

$$\Phi(f)(g, z) := f(g, 0, z) \quad \forall f \in \mathcal{A}_2 \tag{3.372}$$

defines a homomorphism of $(\mathcal{A}_2, \sigma_t)$ to the quantum statistical mechanical system $(\mathcal{B}_2, \sigma_t)$ which is obtained by specialization of the algebra to $\rho = 0$, with convolution product

$$f_1 * f_2(g, z) = \sum_{h \in \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})} f_1(gh^{-1}, h(z)) f_2(h, z). \tag{3.373}$$

Notice that here there is no restriction on the summation, like in the case of the algebras considered in [**93**].
Consider the isometries $\mu_{[n]}$ defined in (3.238) above, and their image

$$\nu_{[n]} := \Phi(\mu_{[n]}). \tag{3.374}$$

The $\nu_{[n]}$ are unitary multipliers of the algebra $\mathcal{B}_2$ and are eigenvectors of the time evolution

$$\sigma_t(\nu_{[n]}) = n^{2it}\nu_{[n]} \tag{3.375}$$

since the $\mu_{[n]}$ are eigenvectors as in (3.242).
The KMS condition shows that, in the presence of unitaries that are eigenvectors of the time evolution as in (3.375), the system $(\mathcal{B}_2, \sigma_t)$ cannot have any KMS_β state. In fact, any such functional would satisfy

$$1 = \varphi(\nu_{[n]} * \nu_{[n]}^*) = \varphi(\nu_{[n]}^* * \sigma_{i\beta}(\nu_{[n]})) = n^{-2\beta}$$

which is impossible for $\beta > 0$. This shows that the pushforward of a KMS state φ of $(\mathcal{A}_2, \sigma_t)$ under the map Φ vanishes identically.
Note that the above proof does not really depend upon the use of multipliers. One can proceed directly and consider for any function $t \in C_c(\Gamma\backslash\mathbb{H})$ and integer $n \neq 1$ the element s of $\mathcal{B}_2$ which vanishes for $g \notin [n]\Gamma$ and fulfills $s([n], z) = t(z)$ for $z \in \mathbb{H}$. The above argument applied using the KMS_β condition for s and s^* shows that φ vanishes on the subalgebra $C_c(\Gamma\backslash\mathbb{H}) \subset \mathcal{B}_2$ which gives the result since this subalgebra contains an approximate unit for $\mathcal{B}_2$.
Recall the discussion on pushforward of KMS states in §2.5 above. Let $C = p_\Gamma(\{(0, z)| z \in \mathbb{H}\})$ be the subset of $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ as in Lemma 3.94. Consider functions $h_n \in C_c(X)$ with values $0 \le h_n(x) \le 1$ and with support disjoint from the set C. Assume that the h_n converge pointwise

$$\lim_{n\to\infty} h_n(\rho, z) = 1$$

in the complement of the set C. Consider then the elements $u_n \in \mathcal{A}_2$ defined by

$$u_n(1, \rho, z) := h_n \circ p_\Gamma(\rho, z) \quad \text{and} \quad u_n(g, \rho, z) = 0, \ \ \forall g \notin \Gamma. \tag{3.376}$$

By Proposition 3.16, we obtain

$$\varphi(f) = \lim_n \varphi(f * u_n). \tag{3.377}$$

Thus, to prove that $\varphi(f) = 0$, we can replace f by $f * u_n$ and since $(f * u_n)(g, y) = f(g, y)\, h_n \circ p_\Gamma(y)$, this means that we can assume that $f(g, y) = 0$ unless $p(y) \in K$, where $K \subset \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ is a compact subset disjoint from the set C. Let $x \in K$ and V as in Lemma 3.94 and let $h \in C_c(V)$. We have

$$h \circ p_\Gamma(g\, y)\ h \circ p_\Gamma(y) = 0\,, \quad \forall y \in Y \tag{3.378}$$

by Lemma 3.94. We then apply the KMS_β condition to the pair a, b with $a = f$ and

$$b(1, \rho, z) := h \circ p_\Gamma(\rho, z) \quad \text{and} \quad b(g, \rho, z) = 0, \ \ \forall g \notin \Gamma. \tag{3.379}$$

By (3.379) we have $\sigma_t(b) = b$ hence the KMS condition gives

$$\varphi(b * f) = \varphi(f * \sigma_{i\beta}(b)) = \varphi(f * b).$$

One has $(b * f)(g, y) = h \circ p_\Gamma(gy)\, f(g, y)$. Applying this to $f * b$ instead of f and using (3.378), we get $\varphi(f * b^2) = 0$. This gives $\varphi(f) = 0$, using a partition of unity on K. □

In the case of the BC system the relevant information to perform the classification of KMS states is obtained from averaging such states under the compact symmetry group G. Equivalently this amounts to studying the restriction of such states to the fixed subalgebra under G. For the system $(\mathcal{A}_2, \sigma_t)$ there is a similar subsystem that will play an important role in the classification. It is obtained by considering functions which are invariant under right multiplication by $\mathrm{GL}_2(\hat{\mathbb{Z}})$ and are independent of the variable $z \in \mathbb{H}$.

Consider the determinant map $\det : M_2(\hat{\mathbb{Z}}) \to \hat{\mathbb{Z}}$, namely the continuous map

$$M_2(\hat{\mathbb{Z}}) \ni \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc \in \hat{\mathbb{Z}}. \tag{3.380}$$

For each $n \in \mathbb{N} = \mathbb{Z}_{>0}$, the composition $\pi_n \circ \det$ defines a projection $\pi'(n)$, which is the characteristic function of the set of $\mathbb{Q}$-lattices whose determinant is divisible by n. Following (3.241) we denote by $\pi(n)$ the projection associated to the characteristic function of the set of $\mathbb{Q}$-lattices which are divisible by n, i.e. with $[n]^{-1}\rho \in M_2(\hat{\mathbb{Z}})$. If a $\mathbb{Q}$-lattice is divisible by n, then its determinant is divisible by n^2. One controls divisibility using the following family of projections $\pi_p(k,l)$. Given a prime p and a pair (k,l) of integers $k \le l$, we let

$$\pi_p(k,l) := (\pi(p^k) - \pi(p^{k+1}))(\pi'(p^{k+l}) - \pi'(p^{k+l+1})). \tag{3.381}$$

This corresponds, when working modulo $N = p^b, b > l$, to matrices in the double class of

$$\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \quad v_p(a) = k \quad \text{and} \qquad v_p(d) = l, \tag{3.382}$$

where v_p is the p-adic valuation.

LEMMA 3.96. (1) *Let φ be a KMS$_\beta$ state for the system $(\mathcal{A}_2, \sigma_t)$. Then, for any prime p and any pair (k,l) of integers $k < l$, one has*

$$\varphi(\pi_p(k,l)) = p^{-(k+l)\beta}\, p^{l-k}\, (1 + p^{-1})\, (1 - p^{-\beta})\, (1 - p^{1-\beta}),$$

while for $k = l$ one has

$$\varphi(\pi_p(l,l)) = p^{-2l\beta}\, (1 - p^{-\beta})\, (1 - p^{1-\beta}).$$

(2) *For distinct primes p_j one has*

$$\varphi(\prod \pi_{p_j}(k_j, l_j)) = \prod \varphi(\pi_{p_j}(k_j, l_j)).$$

PROOF. (1) For each $n \in \mathbb{N}$ we let $\nu_n \in M(A)$ be given by

$$\nu_n(g, \rho, z) = 1 \quad \forall g \in \Gamma \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} \Gamma \quad \text{and} \quad \nu_n(g, \rho, z) = 0 \ \text{ otherwise.}$$

One has $\sigma_t(\nu_n) = n^{it}\, \nu_n$ for all $t \in \mathbb{R}$. The double class

$$\Gamma \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} \Gamma$$

is the union of the left Γ-cosets of the matrices $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ where $ad = n$ and $\gcd\{a, b, d\} = 1$. The number of these left cosets is

$$\omega(n) := n \prod_{p \,\text{prime},\, p|n} (1 + p^{-1})$$

and

$$\nu_n^* * \nu_n(1,\rho,z) = \omega(n), \quad \forall(\rho,z) \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}. \tag{3.383}$$

One has

$$\nu_n * \nu_n^*(1,\rho,z) = \sum_{h \in \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q}),\, h\rho \in M_2(\hat{\mathbb{Z}})} \nu_n(h^{-1}, h\rho, h(z))^2.$$

The right hand side is independent of $z \in \mathbb{H}$ and only depends upon the $\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})$–$\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ double class of $\rho_n = p_n(\rho) \in M_2(\mathbb{Z}/n\mathbb{Z})$.
Let us assume that $n = p^l$ is a prime power. We can assume that $\rho_n = p_n(\rho)$ is of the form

$$\rho_n = \begin{pmatrix} p^a & 0 \\ 0 & p^b \end{pmatrix}, \quad 0 \le a \le b \le l\,.$$

We need to count the number $\omega(a,b)$ of left Γ-cosets Γh_j in the double class

$$\Gamma \begin{pmatrix} n^{-1} & 0 \\ 0 & 1 \end{pmatrix} \Gamma$$

such that $h_j\rho \in M_2(\hat{\mathbb{Z}})$. A full set of representatives of the double class is given by $h_j = (\alpha_j^t)^{-1}$, where the α_j are

$$\alpha_0 = \begin{pmatrix} n & 0 \\ 0 & 1 \end{pmatrix} \quad \alpha(s) = \begin{pmatrix} 1 & s \\ 0 & n \end{pmatrix}, \quad s \in \{0,1,\dots,n-1\}$$

and for $x \in \{1,2,\dots,l-1\}$, $s \in \mathbb{Z}/p^{l-x}\mathbb{Z}$ prime to p

$$\alpha(x,s) = \begin{pmatrix} p^x & s \\ 0 & p^{l-x} \end{pmatrix}.$$

The counting gives

- $\omega(a,b) = 0$ if $b < l$.
- $\omega(a,b) = p^a$ if $a < l, b \ge l$.
- $\omega(a,b) = p^l(1+p^{-1})$ if $a \ge l$.

Let $e_p(i,j)$ for $i \le j$ be the projection corresponding to $a \ge i$, $b \ge j$. Then for $i < j$ one has

$$\pi_p(i,j) = e_p(i,j) - e_p(i+1,j) - e_p(i,j+1) + e_p(i+1,j+1) \tag{3.384}$$

while

$$\pi_p(j,j) = e_p(j,j) - e_p(j,j+1). \tag{3.385}$$

The computation above gives

$$\nu_n * \nu_n^*(1,\rho,z) = p^l(1+p^{-1})e_p(l,l) + \sum_0^{l-1} p^k(e_p(k,l) - e_p(k+1,l)), \tag{3.386}$$

where, for simplicity of notation, we omit the variables (ρ,z) on the right hand side.
Let φ be a KMS_β state, and $\sigma(k,l) := \varphi(e_p(k,l))$. Then, applying the KMS_β condition to the pair $(\mu_{[p]}f, \mu_{[p]}^*)$ for $f \in C(\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H}))$, one gets

$$\sigma(k,l) = p^{-2k\beta}\sigma(0,l-k).$$

Let $\sigma(k) = \sigma(0,k)$. Upon applying the KMS$_\beta$ condition to (ν_n, ν_n^*), one gets

$$p^l(1+p^{-1})p^{-l\beta} = p^l(1+p^{-1})p^{-2l\beta} + \sum_{k=0}^{l-1} p^k(p^{-2k\beta}\sigma(l-k) - p^{-2(k+1)\beta}\sigma(l-k-1).$$

Since $\sigma(0) = 1$, this determines the $\sigma(n)$ by induction on n and gives

$$\sigma(n) = a\, p^{n(1-\beta)} + (1-a)\, p^{-2n\beta},$$

with

$$a = (1+p)\frac{p^\beta - 1}{p^{1+\beta} - 1}\,.$$

Combined with (3.384) and (3.385), this gives the required formulas for $\varphi(\pi_p(k,l))$ and the first part of the Lemma follows.

(2) To get the second part, one proceeds by induction on the number m of primes p_j. The function $f = \prod_1^{m-1} \pi_{p_j}(k_j, l_j)$ fulfills

$$f(h(\rho,z)) = f(\rho,z), \quad \forall(\rho,z) \in M_2(\hat{\mathbb{Z}}) \times \mathbb{H}, \quad \text{and} \quad \forall h \in \Gamma \begin{pmatrix} n^{-1} & 0 \\ 0 & 1 \end{pmatrix} \Gamma,$$

where $n = p_m^l$. Thus, when applying the KMS$_\beta$ condition to $(\nu_n\, f, \nu_n^*)$, the previous computation applies with no change to give the result. $\square$

The main result of this section is the complete classification of low-temperature KMS states, which is obtained as follows.

THEOREM 3.97. *Consider the quantum statistical mechanical system* $(\mathcal{A}_2, \sigma_t)$*.*

(1) *Each invertible* $\mathbb{Q}$*-lattice* $L = (\Lambda, \phi)$ *determines a positive energy representation* π_L *of* $(\mathcal{A}_2, \sigma_t)$*.*
(2) *The partition function of the system is*

$$Z(\beta) = \zeta(\beta)\zeta(\beta - 1) \tag{3.387}$$

(3) *For* $\beta > 2$*, invertible* $\mathbb{Q}$*-lattices* $L = (\Lambda, \phi)$ *determine corresponding extremal KMS*$_\beta$ *states of the form*

$$\varphi_{\beta,L}(f) = Z(\beta)^{-1} \sum_{m \in \Gamma\backslash M_2^+(\mathbb{Z})} f(1, m\rho, m(z)) \det(m)^{-\beta}. \tag{3.388}$$

(4) *For all* $\beta > 2$ *the map* $L \mapsto \varphi_{\beta,L}$ *as in* (3.388) *above gives an identification*

$$\mathcal{E}_\beta = \mathrm{GL}_2^+(\mathbb{Q})\backslash(\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}) = \mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}})/\mathbb{C}^* \tag{3.389}$$

between the set of extremal KMS$_\beta$ *states and the set of invertible* $\mathbb{Q}$*-lattices up to scaling.*

PROOF. (1) We have seen that a $\mathbb{Q}$-lattice $L = (\Lambda, \phi)$ determines a representations π_y of $\mathcal{A}_2$ on the Hilbert space $\mathcal{H}_y = \ell^2(\Gamma\backslash G_y)$ as in (3.202), where $y = (\rho, z)$ is the element in $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ corresponding to the $\mathbb{Q}$-lattice L. In this representation, the time evolution is implemented by the diagonal operator H_y of the form

$$(H_y\, \xi)(g) = \log\det(g)\ \xi(g). \tag{3.390}$$

Namely, we have

$$\pi_y(\sigma_t(f)) = e^{itH_y}\pi_y(f)e^{-itH_y}, \quad \forall f \in \mathcal{A}_2. \tag{3.391}$$

In general, the operator H_y is not positive. However, if the $\mathbb{Q}$-lattice $L = (\Lambda, \phi)$ is invertible, i.e. if $\rho \in \mathrm{GL}_2(\hat{\mathbb{Z}})$, then $G_y = M_2^+(\mathbb{Z})$. In fact, in this case the basis elements of the Hilbert space $\mathcal{H}_y$ are labeled by the lattices $\Lambda' \supset \Lambda$ (by Lemma 3.93). Equivalently, the basis elements are labeled by the elements in the coset space $\Gamma\backslash M_2^+(\mathbb{Z})$. The eigenvalues of H_y are then given by

$$\{\log\det(g)\,, \quad g \in G_y = M_2^+(\mathbb{Z})\}\,,$$

which shows that H_y has positive spectrum.

(2) By (1) the Hilbert space and the Hamiltonian are the same for all the positive energy representations associated to invertible $\mathbb{Q}$-lattices. The dimension of the n-th eigenspace of the Hamiltonian is given by the cardinality of the sets $\{m \in M_2^+(\mathbb{Z}) : \det(m) = n\}$ modulo the action of Γ. Any sublattice of $\mathbb{Z}^2$ is uniquely of the form

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mathbb{Z}^2 \tag{3.392}$$

with $a, d \geq 1$ and $0 \leq b < d$ (cf. [**262**] p.161). Thus, the dimension of the n-th eigenspace is given by the cardinality of the set of matrices as in (3.392) with determinant n. This is given by the formula

$$\sigma_1(n) = \sum_{d|n} d. \tag{3.393}$$

Thus, one obtains that

$$Z(\beta) = \mathrm{Tr}(e^{-\beta H}) = \sum_{n=1}^{\infty} \frac{\sigma_1(n)}{n^\beta} = \zeta(\beta)\zeta(\beta - 1), \tag{3.394}$$

where $\zeta(\beta)$ is the Riemann zeta function.

(3) The state (3.388) is of the form

$$\varphi_{\beta,L}(f) = \frac{\mathrm{Tr}(\pi_L(f)e^{-\beta H})}{\mathrm{Tr}(e^{-\beta H})}$$

where π_L is the positive energy representation associated to the invertible $\mathbb{Q}$-lattice L. Thus, it satisfies the KMS condition at inverse temperature β for $(\mathcal{A}_2, \sigma_t)$.

(4) Let φ be a KMS_β state. Theorem 3.95 shows that there is a probability measure μ on $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$ such that

$$\varphi(f) = \int_{\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H})} f(1, \rho, z)\, d\mu(\rho, z) \quad \forall f \in \mathcal{A}_2.$$

Lemma 3.96 shows that the probability $\varphi(e_p(1,1)) = \sigma(1,1)$ that a prime p divides ρ is $p^{-2\beta}$. Since the series $\sum p^{-2\beta}$ converges ($\beta > \frac{1}{2}$ would suffice here), it follows (cf. [**257**] Thm. 1.41) that, for almost all $(\rho, z) \in \Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$, ρ is only divisible by a finite number of primes. Next, again by Lemma 3.96, the probability that the determinant of ρ is divisible by p is

$$\varphi(e_p(0,1)) = \sigma(1) = (1+p)p^{-\beta} - p^{1-2\beta}.$$

For $\beta > 2$ the corresponding series $\sum ((1+p)p^{-\beta} - p^{1-2\beta})$ is convergent. Thus, we conclude that with probability one

$$\rho_p \in \mathrm{GL}_2(\mathbb{Z}_p) \quad \text{for almost all } p.$$

Moreover, since $\sum \varphi(\pi_p(k,l)) = 1$, one gets with probability one

$$\rho_p \in \mathrm{GL}_2(\mathbb{Q}_p) \quad \forall p.$$

In other words, the measure μ gives measure one to finite idèles (note that finite idèles form a Borel subset but not a closed subset). When ρ is a finite idèle the corresponding $\mathbb{Q}$-lattice is commensurable to a unique invertible $\mathbb{Q}$-lattice. Then the KMS_β condition shows that the measure μ is entirely determined by its restriction to invertible $\mathbb{Q}$-lattices, so that, for some probability measure ν, we have

$$\varphi = \int \varphi_{\beta,L}\, d\nu(L).$$

It follows that the Choquet simplex of extremal KMS_β states is the space of probability measures on the locally compact space

$$\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})/\mathbb{C}^* \tag{3.395}$$

of invertible $\mathbb{Q}$-lattices and its extreme points are the $\varphi_{\beta,L}$. See e.g. [**234**] for the standard identification of the set of invertible $\mathbb{Q}$-lattices with the double quotient (3.395). ☐

Proposition 3.98. *The set $\mathcal{E}_\infty$ of extremal KMS_∞ states of the system $(\mathcal{A}_2, \sigma_t)$ is identified with the set of invertible 2-dimensional $\mathbb{Q}$-lattices up to scaling, that is, with the adelic quotient* (3.395)*.*

Proof. In the limit $\beta \to \infty$ the states $\varphi_{\beta,L}$ of (3.388) converge and their limits restrict to

$$C_c(\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})) \subset A$$

as characters given by evaluation

$$\varphi_{\infty,L}(f) = f(\rho, z), \quad \forall f \in C_c(\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})), \tag{3.396}$$

where the invertible $\mathbb{Q}$-lattice L is identified up to scaling with the data $(\rho, z) \in \mathrm{GL}_2(\hat{\mathbb{Z}}) \times_\Gamma \mathbb{H}$.

The characters (3.396) are all distinct, hence we get a bijection of the space

$$\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})/\mathbb{C}^*$$

of invertible $\mathbb{Q}$-lattices with the space $\mathcal{E}_\infty$ of extremal KMS_∞ states. ☐

7.5. Action of symmetries on KMS states.

We can now describe explicitly the action of the symmetry group $S = \mathbb{Q}^*\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ on the low-temperature KMS_β states, with $\beta > 2$.

The following result describes the induced action of the endomorphisms ϑ_m of (3.232), for $m \in M_2(\hat{\mathbb{Z}})$, on the extremal low-temperature KMS states.

PROPOSITION 3.99. *For a given $\beta > 2$, let $\varphi_{\beta,L}$ be the extremal KMS$_\beta$ state* (3.388) *associated to the invertible 2-dimensional $\mathbb{Q}$-lattice $L = (\Lambda, \phi)$ specified up to scale by the element $(\rho, z) \in \mathrm{GL}_2(\hat{\mathbb{Z}}) \times_\Gamma \mathbb{H}$. For $m \in M_2^+(\mathbb{Z})$, the pullback of the state $\varphi_{\beta,L}$ under the action of ϑ_m is of the form*

$$\varphi_{\beta,L} \circ \vartheta_m = \varphi_{\beta,L}(e_{m\mathbb{Z}^2})\, \varphi_{\beta,L'}, \qquad \text{with } L' = (\rho', m'^{-1}(z)), \tag{3.397}$$

where $\rho m = m'\rho'$ is the decomposition in $M_2^+(\mathbb{Z}) \cdot \mathrm{GL}_2(\hat{\mathbb{Z}})$ of the product $\rho m \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ and $e_{m\mathbb{Z}^2}$ are as in (3.234).

PROOF. One has $\rho m \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) = \mathrm{GL}_2^+(\mathbb{Q}) \cdot \mathrm{GL}_2(\hat{\mathbb{Z}})$, hence one can write $\rho m = m'\rho'$, with $\rho' \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ and $m' \in \mathrm{GL}_2^+(\mathbb{Q})$. In fact, one has $m' \in M_2(\hat{\mathbb{Z}})$, since $\rho \in \mathrm{GL}_2(\hat{\mathbb{Z}})$. Since $M_2(\hat{\mathbb{Z}}) \cap \mathrm{GL}_2^+(\mathbb{Q}) = M_2^+(\mathbb{Z})$, one obtains that $m' \in M_2^+(\mathbb{Z})$. The decomposition is not unique and one can replace (ρ', m') by $(\gamma\rho', m'\gamma^{-1})$ for $\gamma \in \Gamma$. But this replacement does not change the $\mathbb{Q}$-lattice $L' = (\Lambda', \phi')$ specified by $(\rho', m'^{-1}(z))$. Notice also that $\det(\rho m) = \det(m'\rho')$. This gives

$$\det m = \det m', \qquad \text{and} \qquad \det\rho = \det\rho'. \tag{3.398}$$

Thus we obtain

$$\rho m \det(m)^{-1} = m' \det(m')^{-1}\rho' = (\tilde{m}')^{-1}\rho' \tag{3.399}$$

with $\tilde{m}'$ as in (3.235). This gives

$$\varphi_{\beta,L}(\vartheta_m(f)) = Z_\beta^{-1} \sum_{\mu \in \Gamma\backslash M_2^+(\mathbb{Z})} f(1, \mu\rho m \det(m)^{-1}, \mu m)(\det\mu)^{-\beta} \tag{3.400}$$

where $\mu \in M_2^+(\mathbb{Z})$ is subject to the condition $\mu\rho m \det(m)^{-1} \in M_2(\hat{\mathbb{Z}})$. The other values of μ a priori involved in the summation do not contribute, since they correspond to the orthogonal complement of the support $\subset e_{m\mathbb{Z}^2}$ of $\vartheta_m(f)$. By (3.399) one has $\rho m \det(m)^{-1} = (\tilde{m}')^{-1}\rho'$. Therefore, since $\rho' \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ the condition $\mu\rho m \det(m)^{-1} \in M_2(\hat{\mathbb{Z}})$ holds iff $\mu(\tilde{m}')^{-1} \in M_2(\hat{\mathbb{Z}})$ i.e. iff $\mu = \nu\tilde{m}'$ for some $\nu \in M_2^+(\mathbb{Z})$ (since $\mathrm{GL}_2^+(\mathbb{Q}) \cap M_2(\hat{\mathbb{Z}}) = M_2^+(\mathbb{Z})$).
Thus, one can replace the sum in (3.400) by

$$\begin{aligned} &\varphi_{\beta,L}(\vartheta_m(f)) = \\ &Z_\beta^{-1} \sum_{\nu \in \Gamma\backslash M_2^+(\mathbb{Z})} f(1, \nu\, \underbrace{\tilde{m}'\rho m \det(m)^{-1}}_{\rho'}, \nu\tilde{m}'(z)) \det(\nu\tilde{m}')^{-\beta}. \end{aligned} \tag{3.401}$$

This is proportional to $\varphi_{\beta,L'}$, since $\tilde{m}'(z) = m'^{-1}(z)$. □

7.6. Low-temperature KMS states and Galois action.

We can now state the main result on the arithmetic properties of the system $(\mathcal{A}_2, \sigma_t)$.

THEOREM 3.100. *Let $L = (\Lambda, \phi)$ be a generic invertible $\mathbb{Q}$-lattice, that is, an invertible $\mathbb{Q}$-lattice represented by an element (ρ, τ) with $\rho \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ and $\tau \in \mathbb{H}$ a generic point (i.e. $j(\tau)$ transcendental). Let $\varphi_{\infty,L} \in \mathcal{E}_\infty$ be the corresponding extremal KMS$_\infty$ state on $(\mathcal{A}_2, \sigma_t)$. The image $\varphi_{\infty,L}(\mathcal{A}_{2,\mathbb{Q}}) \subset \mathbb{C}$ generates the specialization $F_\tau \subset \mathbb{C}$ of the modular field F at the point $\tau \in \mathbb{H}$. The action of the* *symmetry group* $S = \mathbb{Q}^*\backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ *of the dynamical system* $(\mathcal{A}_2, \sigma_t)$ *is intertwined by* $\varphi_{\infty,L}$ *with the Galois group of the modular field* F_τ *by the formula*

$$\varphi_{\infty,L} \circ \vartheta_\alpha(f) = \mathrm{Gal}_\tau(\rho\alpha\rho^{-1})\, \varphi_{\infty,L}(f), \tag{3.402}$$

for all $\alpha \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ *and all* $f \in \mathcal{A}_{2,\mathbb{Q}}$.

Proof. One has

$$\varphi_{\infty,L}(f) = f(1,\rho,\tau) \quad \forall f \in \mathcal{A}_{2,\mathbb{Q}}, \tag{3.403}$$

where the invertible $\mathbb{Q}$-lattice $L = (\Lambda, \phi)$ is represented by the point $(\rho,\tau) \in \mathrm{GL}_2(\hat{\mathbb{Z}}) \times \mathbb{H}$.
By Definition 3.83 one has $\varphi_{\infty,L}(\mathcal{A}_{2,\mathbb{Q}}) \subset F_\tau$. Let $\mathcal{X}_a \in \mathcal{A}_{2,\mathbb{Q}}$ be as in §7.1. Up to a permutation of the labels the $\varphi_{\infty,L}(\mathcal{X}_a)$ are the $\mathfrak{f}_u(\tau)$ and thus generate F_τ. As we have seen in §6, it already contains $\mathbb{Q}(j)$ at level 2 and it in fact contains $\mathbb{Q}^{\mathrm{ab}}(j)$.
Let us now show the intertwining equality (3.402). It is enough to prove (3.402) for $\alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})$ and for $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$.
For $\alpha \in \mathrm{GL}_2(\hat{\mathbb{Z}})$, the state $\varphi_{\infty,L} \circ \vartheta_\alpha$ is given simply by (3.219) i.e.

$$(\varphi_{\infty,L} \circ \vartheta_\alpha)(f) = f(1,\rho\alpha,\tau), \qquad \forall f \in \mathcal{A}_{2,\mathbb{Q}},$$

and for $L = (\rho,\tau)$ as above. Using the property (3.333) of elements $f \in \mathcal{A}_{2,\mathbb{Q}}$ one gets (3.402) in this case since

$$f(1,\rho\alpha,\tau) = \mathrm{Gal}_\tau(\rho\alpha\rho^{-1})\, f(1,\rho,\tau) = \mathrm{Gal}_\tau(\rho\alpha\rho^{-1})\, \varphi_{\infty,L}(f)\,.$$

Consider then the case of $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$. In fact one can assume by Theorem 3.56 that α is an element $m \in M_2^+(\mathbb{Z})$. The state $\varphi_{\infty,L} \circ m$ is more tricky to obtain. In fact, it is not given by straight composition, but through the warming up/cooling down process described in §2.4. Indeed, the range of the transformation ϑ_m of (3.232) is the reduced algebra by the projection $e_{m(\mathbb{Z}^2)}$, on which any of the zero-temperature states vanishes identically. This makes it impossible to define the pullback $\vartheta_m^*(\varphi_{\infty,L})$ as in (3.23). A non-trivial action is obtained by taking the weak limit of the KMS_β states obtained by the action of ϑ_m on the KMS_β state $\varphi_{\beta,L}$,

$$\vartheta_m^*(\varphi_{\infty,L}) = \lim_{\beta\to\infty} \vartheta_m^*(\varphi_{\beta,L}),$$

where $\vartheta_m^*(\varphi_{\beta,L}) = \varphi_{\beta,L}(e_{m\mathbb{Z}^2})^{-1}\varphi_{\beta,L} \circ \vartheta_m$ as in (3.23).
By Proposition 3.99 we have for finite β

$$\varphi_{\beta,L} \circ \vartheta_m = \varphi_{\beta,L}(e_{m\mathbb{Z}^2})\, \varphi_{\beta,L'}, \tag{3.404}$$

where L' is given by

$$L' = (\rho', m'^{-1}(\tau)), \qquad \text{with} \ \ \rho m = m'\rho' \in M_2^+(\mathbb{Z})\mathrm{GL}_2(\hat{\mathbb{Z}}). \tag{3.405}$$

Thus, we get

$$\vartheta_m^*(\varphi_{\infty,L}) = \varphi_{\infty,L'}\,. \tag{3.406}$$

It remains to show that on $\mathcal{A}_{2,\mathbb{Q}}$ we have

$$\varphi_{\infty,L'}(f) = \mathrm{Gal}_\tau(\rho m\rho^{-1}) \circ \varphi_{\infty,L}(f), \qquad \forall f \in \mathcal{A}_{2,\mathbb{Q}}.$$

Both sides only involve the values of f on invertible $\mathbb{Q}$-lattices. There, by the property (3.333) of elements of $\mathcal{A}_{2,\mathbb{Q}}$ one has

$$f(1,\alpha,\tau) = I_\tau(\mathrm{Gal}(\alpha)f_{(1,1)}) = \mathrm{Gal}_\tau(\alpha)I_\tau(f_{(1,1)})\,, \qquad \forall \alpha \in \mathrm{GL}(2,\hat{\mathbb{Z}}).$$

Thus, if we let $h = f_{(1,1)} \in F$ we obtain

$$\varphi_{\infty,L'}(f) = f(1,\rho', m'^{-1}(\tau)) = I_{m'^{-1}(\tau)}(\mathrm{Gal}(\rho')h) = I_\tau(\mathrm{Gal}(m'\rho')h).$$

using (3.340) to get

$$I_{g^{-1}(\tau)} = I_\tau \circ \mathrm{Gal}(g)$$

for $g = m'$. Since $m'\rho' = \rho m$, we obtain

$$I_\tau(\mathrm{Gal}(m'\rho')h) = \mathrm{Gal}_\tau(\rho m \rho^{-1})\mathrm{Gal}_\tau(\rho) I_\tau(h) = \mathrm{Gal}_\tau(\rho m \rho^{-1})\varphi_{\infty,L}(f)$$

as required. □

7.7. The high temperature range.

The fact that the partition function $Z(\beta) = \zeta(\beta)\zeta(\beta-1)$ of the system $(\mathcal{A}_2, \sigma_t)$ has poles at $\beta = 1$ and $\beta = 2$ suggests the presence of a more complicated pattern of phase transitions than in the case of the BC system $\mathcal{A}_1$. In fact, we encounter a new phenomenon, which did not occur in the case of 1-dimensional $\mathbb{Q}$-lattices, namely, at sufficiently high temperatures there are no KMS states at all (cf. [**86**]).

PROPOSITION 3.101. *The system $(\mathcal{A}_2, \sigma_t)$ has no KMS$_\beta$ state for any $\beta < 1$.*

PROOF. The result follows directly from Lemma 3.96. Indeed the value of $\varphi(\pi_p(k,l))$ provided by the lemma is strictly negative for $\beta < 1$. □

REMARK 3.102. It was wrongly asserted in [**86**] that there is no KMS$_\beta$ state for $\beta = 1$. While it is true that the values $\varphi(\pi_p(k,l))$ all vanish for $\beta = 1$, this does not suffice to conclude that the measure is carried by the subset $\{0\} \times \mathbb{H}$ and in fact, it was shown in [**193**] that there are KMS$_\beta$ states for $\beta = 1$ with support on singular matrices.

The existence of KMS$_\beta$ states in the range $\beta \in (1,2]$ and the uniqueness for $\beta = 2$ were suggested in [**86**] Lemma 8.13, and a definitive result on the KMS states of the system $(\mathcal{A}_2, \sigma_t)$ in the range $\beta \in (1,2]$ was obtained recently by Laca–Larsen–Neshveyev [**193**]. They show the following uniqueness result.

Let $\beta > 1$ and $\nu_{\beta,p}$ be the Haar measure on $\mathrm{GL}_2(\mathbb{Z}_p)$ normalized to give total mass

$$\nu_{\beta,p}(\mathrm{GL}_2(\mathbb{Z}_p)) = (1-p^{-\beta})(1-p^{-\beta+1}) \tag{3.407}$$

and let $\mu_{\beta,p}$ be the unique extension of $\nu_{\beta,p}$ to $\mathrm{GL}_2(\mathbb{Q}_p)$ such that for all $g \in \mathrm{GL}_2(\mathbb{Q}_p)$ one has,

$$\mu_{\beta,p}(g\mathcal{K}) = |\det(g)|_p^\beta \, \mu_{\beta,p}(\mathcal{K}) \quad \text{ for compact } \mathcal{K} \subset \mathrm{GL}_2(\mathbb{Q}_p)\,. \tag{3.408}$$

THEOREM 3.103. ([**193**]) *For any $\beta \in (1,2]$, the set of KMS$_\beta$ states of the system $(\mathcal{A}_2, \sigma_t)$ consists of a single state φ_β. This is given by a Γ-invariant measure μ_β on $M_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}$ satisfying*

$$\mu_\beta(\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})) = 1 \quad \text{ and } \quad \mu_\beta(g\mathcal{K}) = \det(g)^{-\beta}\mu(\mathcal{K}),$$

for all $g \in \mathrm{GL}_2^+(\mathbb{Q})$ and all compact $\mathcal{K} \subset M_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}$. The measure is of the form $\mu_\beta = \mu_{\beta,f} \times \mu_{\mathbb{H}}$, with $\mu_{\mathbb{H}}$ as in (3.207) *and $\mu_{\beta,f} = \prod_p \mu_{\beta,p}$.*

Note that the above normalization (3.407) no longer makes sense when $\beta \leq 1$. Theorem 3.103 allows one to improve the result obtained in [**86**] Lemma 8.13 on the convergence of the KMS_β states $\varphi_{\beta,L}$ when $\beta \to 2+$, independently of the invertible $\mathbb{Q}$-lattice L. This convergence is now seen to hold on all continuous functions $f \in C(\Gamma\backslash(M_2(\hat{\mathbb{Z}})\times\mathbb{H}))$. Another interesting result of [**193**] Remark 4.10 is the description of the classification of KMS_β states for n-dimensional $\mathbb{Q}$-lattices modulo scaling by $\mathbb{R}_+^*$. One gets that for $\beta \leq n-1$ there are no KMS_β states except when β is a positive integer. Next in the interval $\beta \in (n-1, n]$ there exists a unique KMS_β state, and for $\beta > n$ the extremal KMS_β states are parameterized by invertible $\mathbb{Q}$-lattices modulo scaling.

7.8. The Shimura variety of GL_2.

We have seen in Theorem 3.97 and Proposition 3.98 that the set of extremal KMS states for any $2 < \beta \leq \infty$ is identified with the adelic quotient (3.395). This identification, together with the description of the set of two-dimensional $\mathbb{Q}$-lattices up to scaling given in Proposition 3.39, yields a description of the system $(\mathcal{A}_2, \sigma_t)$ and its low-temperature KMS equilibrium states in terms of Shimura varieties, as we discussed in the case of the BC system in §4.8 above. We follow [**91**].

We consider here the Shimura variety associated to the data $(\mathrm{GL}_2, \mathbb{H}^\pm)$, where $\mathbb{H}^\pm$ denotes the upper and lower half planes

$$\mathbb{H}^\pm = \mathrm{GL}_2(\mathbb{R})/\mathbb{C}^*. \tag{3.409}$$

Namely, the Shimura variety $Sh(\mathrm{GL}_2, \mathbb{H}^\pm)$ is given by the quotient

$$Sh(\mathrm{GL}_2, \mathbb{H}^\pm) = \mathrm{GL}_2(\mathbb{Q})\backslash(\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}^\pm). \tag{3.410}$$

As in the GL_1-case the zero-dimensional Shimura variety $Sh(\mathrm{GL}_1, \{\pm 1\})$ was associated to the cyclotomic tower, here the Shimura variety (3.410) is associated to the *modular tower*, which here means the tower of modular curves. More precisely, for $N \in \mathbb{N}$, let $Y(N) = \Gamma(N)\backslash\mathbb{H}$ be the modular curve of level N, where $\Gamma(N)$ is the principal congruence subgroup of $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. We denote by $X(N)$ the smooth compactification of $Y(N)$ obtained by adding the cusps $\Gamma(N)\backslash\mathbb{P}^1(\mathbb{Q})$. In the compact version, the modular tower has $V_1 = \mathbb{P}^1$ over $\mathbb{Q}$ and $V_N = X(N)$, over the cyclotomic field $\mathbb{Q}(\zeta_N)$ (cf. [**226**]). In the noncompact version, one has the affine line $V_1 = \mathbb{A}^1$, and $V_N = Y(N)$. Notice that $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\pm 1$ is the group of automorphisms of the projection $V_N = X(N) \to X(1) = V_1$, so that one has

$$\varprojlim_N \mathrm{Aut}_{V_1}(V_N) = \mathrm{GL}_2(\hat{\mathbb{Z}})/\{\pm 1\} = \varprojlim_N \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}. \tag{3.411}$$

This recovers the part of the symmetries of low-temperature KMS states of $(\mathcal{A}_2, \sigma_t)$ that comes from automorphisms of the system. The endomorphism action of $M_2^+(\mathbb{Z})$ does not correspond to symmetries of the tower, in fact, the action of the θ_m does not preserve levels.

The result of Proposition 3.39 shows that, set-theoretically, the commensurability classes of 2-dimensional $\mathbb{Q}$-lattices up to scaling are parameterized by the points of the quotient

$$Sh^{\mathrm{nc}}(\mathrm{GL}_2, \mathbb{H}^\pm) := \mathrm{GL}_2(\mathbb{Q})\backslash(M_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}^\pm). \tag{3.412}$$

At a more refined level than just the set-theoretic identification, the quotient (3.412) is treated as a noncommutative space whose algebra of coordinates is (Morita equivalent to) the convolution algebra $\mathcal{A}_2$, just as in the case of 1-dimensional $\mathbb{Q}$-lattices. The Morita equivalence is realized by compression with the characteristic function of $M_2(\hat{\mathbb{Z}})$ in $M_2(\mathbb{A}_{\mathbb{Q},f})$.

The quotient (3.412) can be regarded as a noncommutative version of the Shimura variety (3.410), obtained by adding to (3.410) "noncommutative boundary strata" that correspond to the commensurability classes of non-invertible $\mathbb{Q}$-lattices.

We can view the classical Shimura variety (3.410) as the set of classical points of the noncommutative space (3.412), through the identification of $Sh(\mathrm{GL}_2, \mathbb{H}^\pm)$ with the set of low-temperature extremal KMS states of $(\mathcal{A}_2, \sigma_t)$ obtained in Theorem 3.97 and Proposition 3.98,

(3.413)
$$\mathcal{E}_\infty = Sh(\mathrm{GL}_2, \mathbb{H}^\pm) = \mathrm{GL}_2(\mathbb{Q})\backslash(\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}^\pm) = \mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}})/\mathbb{C}^*.$$

7.9. The noncommutative boundary of modular curves.

Under the Gelfand–Naimark correspondence, unital C^*-algebras correspond to compact spaces. Thus, we can define a notion of "Morita-compactification" for noncommutative spaces as follows.

DEFINITION 3.104. *A noncommutative space described by a C^*-algebra $\mathcal{A}$ is "Morita-compact" if the algebra $\mathcal{A}$ is Morita equivalent to a unital C^*-algebra.*

Consider then the quotient

$$\overline{Sh^{\mathrm{nc}}(\mathrm{GL}_2, \mathbb{H}^\pm)} = \mathrm{GL}_2(\mathbb{Q})\backslash M_2(\mathbb{A}_{\mathbb{Q}})^{\cdot}/\mathbb{C}^*, \tag{3.414}$$

with

$$M_2(\mathbb{A}_{\mathbb{Q}})^{\cdot} = M_2(\mathbb{A}_{\mathbb{Q},f}) \times (M_2(\mathbb{R}) \smallsetminus \{0\}).$$

LEMMA 3.105. *There is an identification*

$$\mathrm{GL}_2(\mathbb{Q})\backslash M_2(\mathbb{A}_{\mathbb{Q}})^{\cdot}/\mathbb{C}^* = \mathrm{GL}_2(\mathbb{Q})\backslash(M_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{P}^1(\mathbb{C})), \tag{3.415}$$

with $\mathbb{P}^1(\mathbb{C}) = \mathbb{H}^\pm \cup \mathbb{P}^1(\mathbb{R})$ the compactification of $\mathbb{H}^\pm$.

PROOF. Consider the right action

$$m \mapsto m \cdot \Delta \tag{3.416}$$

of $\mathbb{C}^*$ on $M_2(\mathbb{R})$ determined by the inclusion $\mathbb{C}^* \subset \mathrm{GL}_2(\mathbb{R})$ as in (3.170). The action of $\mathbb{C}^*$ on $M_2(\mathbb{R}) \smallsetminus \{0\}$ is free and proper. To obtain the identification consider the map

$$\rho(\alpha) = \begin{cases} \alpha(i) = \frac{ai+b}{ci+d} & (c,d) \neq (0,0) \\ \infty & (c,d) = (0,0) \end{cases} \qquad \text{with } \alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R}) \smallsetminus \{0\}$$

It defines an isomorphism

$$(M_2(\mathbb{R}) \smallsetminus \{0\})/\mathbb{C}^* \xrightarrow{\simeq} \mathbb{P}^1(\mathbb{C}), \tag{3.417}$$

equivariant with respect to the left action of $\mathrm{GL}_2(\mathbb{R})$ on $M_2(\mathbb{R})$ and the action of $\mathrm{GL}_2(\mathbb{R})$ on $\mathbb{P}^1(\mathbb{C})$ by fractional linear transformations. Moreover, this maps $M_2^+(\mathbb{R})$ to the closure of the upper half plane

$$\overline{\mathbb{H}} = \mathbb{H} \cup \mathbb{P}^1(\mathbb{R}). \tag{3.418}$$

The rank one matrices in $M_2(\mathbb{R})$ map to $\mathbb{P}^1(\mathbb{R}) \subset \mathbb{P}^1(\mathbb{C})$. In fact, the isotropy group of $m \in M_2(\mathbb{R})$ for the right action of $\mathbb{C}^*$ is trivial if $m \neq 0$, since $m \cdot \lambda = m$ only has nontrivial solutions for $m = 0$, since $\lambda - 1$ is invertible when nonzero. This shows that $M_2(\mathbb{R}) \smallsetminus \{0\}$ is the total space of a principal $\mathbb{C}^*$-bundle. $\square$

Thus, we can regard (3.414) as a noncommutative space whose algebra of coordinates is Morita equivalent to the convolution algebra of the partially defined action of $\mathrm{GL}_2(\mathbb{Q})$ on the compact space $M_2(\hat{\mathbb{Z}}) \times \mathbb{P}^1(\mathbb{C})$, hence as a Morita compactification, in the sense of Definition 3.104, of the noncommutative Shimura variety (3.412), or equivalently of the algebra $\mathcal{A}_2$.

As we discussed in §7.8 above, the classical compactification of the modular curves $Y(N)$ is given by $X(N) = Y(N) \cup \Gamma(N)\backslash\mathbb{P}^1(\mathbb{Q})$. This is obtained by replacing the upper half plane $\mathbb{H}$ by

$$\bar{\mathbb{H}}_{\mathbb{Q}} := \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q}). \tag{3.419}$$

While the quotient $Y(N) = \Gamma(N)\backslash\mathbb{H}$ parameterizes elliptic curves with level structure, the points $\Gamma(N)\backslash\mathbb{P}^1(\mathbb{Q})$ correspond to degenerating the elliptic curve to $\mathbb{G}_m$. The Morita compactification described above, however, is obtained by replacing $\mathbb{H}$ with the compactification $\bar{\mathbb{H}}$ of (3.418), which uses the full $\mathbb{P}^1(\mathbb{R})$ instead of $\mathbb{P}^1(\mathbb{Q})$ as in (3.419). This corresponds to the noncommutative compactification of modular curves

$$X(N)^{\mathrm{nc}} = \Gamma(N)\backslash\bar{\mathbb{H}} = \Gamma(N)\backslash(\mathbb{H} \cup \mathbb{P}^1(\mathbb{R})) \tag{3.420}$$

studied in [**219**] and [**77**], where the noncommutative space

$$\Gamma(N)\backslash(\mathbb{P}^1(\mathbb{R}) \smallsetminus \mathbb{P}^1(\mathbb{Q})) \tag{3.421}$$

parameterizes degenerations of elliptic curves to noncommutative tori.

In fact, we can give the following description of the space (3.414) in terms of $\mathbb{Q}$-lattices.

DEFINITION 3.106. *A $\mathbb{Q}$-pseudolattice in $\mathbb{C}$ is a pair (Λ, ϕ), where $\Lambda = j(\mathbb{Z}^2)$ is the image of a homomorphism $j : \mathbb{Z}^2 \to \ell$, with $\ell \subset \mathbb{R}^2 \cong \mathbb{C}$ a real 1-dimensional subspace, and with ϕ a group homomorphism*

$$\phi : \mathbb{Q}^2/\mathbb{Z}^2 \to \mathbb{Q}\Lambda/\Lambda.$$

The $\mathbb{Q}$-pseudolattice is nondegenerate if j is injective and is invertible if ϕ is invertible.

Pseudolattices account for replacing $\mathrm{GL}_2(\mathbb{R})$ by matrices $M_2(\mathbb{R}) \smallsetminus \{0\}$, which corresponds to degenerating the lattices to pseudolattices in the sense of [**216**]. In more geometric terms, this corresponds to degenerations of elliptic curves to noncommutative tori.

PROPOSITION 3.107. *Let* $\partial Y := M_2(\hat{\mathbb{Z}}) \times \mathbb{P}^1(\mathbb{R})$. *The map*

$$(\rho, \theta) \mapsto (\Lambda, \phi), \qquad \Lambda = \mathbb{Z} + \theta\mathbb{Z}, \qquad \phi(x) = \rho_1(x) - \theta\rho_2(x) \tag{3.422}$$

gives an identification

$$\Gamma\backslash\partial Y \simeq (\textit{Space of } \mathbb{Q}\textit{-pseudolattices in } \mathbb{C})/\mathbb{C}^*. \tag{3.423}$$

This space parameterizes the degenerations of 2-dimensional $\mathbb{Q}$*-lattices*

$$\lambda(y) = (\Lambda, \phi) \quad \textit{where} \quad \Lambda = \tilde{h}(\mathbb{Z} + i\mathbb{Z}) \quad \textit{and} \quad \phi = \tilde{h} \circ \rho, \tag{3.424}$$

for $y = (\rho, h) \in M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2(\mathbb{R})$ *and* $\tilde{h} = h^{-1}\det(h)$, *when* $h \in \mathrm{GL}_2(\mathbb{R})$ *degenerates to a non-invertible matrix in* $M_2(\mathbb{R}) \smallsetminus \{0\}$.

PROOF. $\mathbb{Q}$-pseudolattices in $\mathbb{C}$ are of the form

$$\Lambda = \lambda(\mathbb{Z} + \theta\mathbb{Z}), \quad \phi(a) = \lambda\rho_1(a) - \lambda\theta\rho_2(a), \tag{3.425}$$

for $\lambda \in \mathbb{C}^*$, $\theta \in \mathbb{P}^1(\mathbb{R})$ and $\rho \in M_2(\hat{\mathbb{Z}})$. The action of $\mathbb{C}^*$ multiplies λ, while leaving θ unchanged. This corresponds to changing the 1-dimensional linear subspace of $\mathbb{C}$ containing the pseudolattice and rescaling it. The action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{R})$ by fractional linear transformations changes θ. The nondegenerate pseudolattices correspond to the values $\theta \in \mathbb{P}^1(\mathbb{R}) \smallsetminus \mathbb{P}^1(\mathbb{Q})$ and the degenerate pseudolattices to the cusps $\mathbb{P}^1(\mathbb{Q})$.
For $y = (\rho, h) \in M_2(\hat{\mathbb{Z}}) \times \mathrm{GL}_2(\mathbb{R})$, consider the $\mathbb{Q}$-lattice (3.424), for $\tilde{h} = h^{-1}\det(h)$. Here we use the basis $\{e_1 = 1, e_2 = -i\}$ of the $\mathbb{R}$-vector space $\mathbb{C}$ so that $\mathrm{GL}_2^+(\mathbb{R})$ acts on $\mathbb{C}$ as $\mathbb{R}$-linear transformations, as in (3.170). These formulae continue to make sense when $h \in M_2(\mathbb{R})$, $h \neq 0$ and the image $\Lambda = \tilde{h}(\mathbb{Z} + i\mathbb{Z})$ is a pseudolattice, i.e. when the matrix h is no longer invertible, cf. Lemma 3.105. ☐

Unlike the quotient $\Gamma\backslash(M_2(\hat{\mathbb{Z}}) \times \mathbb{H})$, which is a classical space, here the space $\Gamma\backslash\partial Y$ should be treated as a noncommutative space, whose algebra of coordinates is given by the crossed product

$$C(\partial Y) \rtimes \Gamma.$$

The "invertible part"

$$\mathrm{GL}_2(\mathbb{Q})\backslash(\mathrm{GL}_2(\mathbb{A}_f) \times \mathbb{P}^1(\mathbb{R})) \tag{3.426}$$

gives the noncommutative modular tower considered in [**77**], [**219**], and [**240**], so that the full space $\overline{Sh^{\mathrm{nc}}(\mathrm{GL}_2, \mathbb{H}^{\pm})}$ accounts for degenerations of invertible $\mathbb{Q}$-lattices to either a non-Archimedean component (non-invertible $\mathbb{Q}$-lattices) or at the Archimedean place (pseudolattices).

7.10. Compatibility between the systems.

Up to Morita equivalence, the noncommutative spaces of $\mathbb{Q}$-lattices modulo commensurability (but not scaling) in dimensions $n = 1, 2$ are described, respectively, by the crossed product C^*-algebras

$$C_0(\mathbb{A}_f \times \mathrm{GL}_1(\mathbb{R})) \rtimes \mathrm{GL}_1(\mathbb{Q}), \tag{3.427}$$

and

$$C_0(M_2(\mathbb{A}_f) \times \mathrm{GL}_2(\mathbb{R})) \rtimes \mathrm{GL}_2(\mathbb{Q}). \tag{3.428}$$

In the case of the classical Shimura varieties, the relation between the varieties $Sh(\mathrm{GL}_2, \mathbb{H}^{\pm})$ and $Sh(\mathrm{GL}_1, \{\pm 1\})$ is given by "passing to components". Namely we have (cf. [**226**])

$$\pi_0(Sh(\mathrm{GL}_2, \mathbb{H}^{\pm})) = Sh(\mathrm{GL}_1, \{\pm 1\}). \tag{3.429}$$

The operation of taking connected components is realized explicitly by the map

$$\mathrm{sign} \times \det : Sh(\mathrm{GL}_2, \mathbb{H}^{\pm}) \to Sh(\mathrm{GL}_1, \{\pm 1\}). \tag{3.430}$$

We showed in [**91**] that the map (3.430) extends to a correspondence between the noncommutative Shimura varieties $Sh^{\mathrm{nc}}(\mathrm{GL}_1, \{\pm 1\})$ and $Sh^{\mathrm{nc}}(\mathrm{GL}_2, \mathbb{H}^{\pm})$, compatible with the time evolutions and the arithmetic structures. The result is obtained as follows (cf. [**91**]).

We consider the map $\alpha : \mathrm{GL}_1(\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Q})$ of the form

$$\alpha(r) = \begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.431}$$

We take π to be the determinant map from $M_2(\mathbb{A}_f) \times \mathrm{GL}_2(\mathbb{R})$ to $\mathbb{A}_f \times \mathrm{GL}_1(\mathbb{R})$ given by:

$$(\rho, u) \mapsto \pi(\rho, u) = (\det(\rho), \det(u)). \tag{3.432}$$

Proposition 3.108. *Let α and π be defined as in* (3.431) *and* (3.432)*. Then the following holds.*

(1) *The pair (π, α) induces a morphism*

$$\Upsilon : C_0(\mathbb{A}_f \times \mathrm{GL}_1(\mathbb{R})) \rtimes \mathrm{GL}_1(\mathbb{Q}) \to M(C_0(M_2(\mathbb{A}_f) \times \mathrm{GL}_2(\mathbb{R}))) \rtimes \mathrm{GL}_2(\mathbb{Q}),$$

where $M(C_0(M_2(\mathbb{A}_f) \times \mathrm{GL}_2(\mathbb{R}))$ is the multiplier algebra of the algebra $C_0(M_2(\mathbb{A}_f) \times \mathrm{GL}_2(\mathbb{R})$.

(2) *The restriction of Υ to homogeneous functions of weight 0 induces a morphism*

$$\Upsilon_0 : C_0(\mathbb{A}_f \times \{\pm 1\}) \rtimes \mathrm{GL}_1(\mathbb{Q}) \to M(C_0(M_2(\mathbb{A}_f) \times \mathbb{H}^{\pm})) \rtimes \mathrm{GL}_2(\mathbb{Q}). \tag{3.433}$$

(3) *The morphism Υ yields a correspondence Υ_{12} from the GL_1-system to the GL_2-system, compatible with the time evolutions and the rational subalgebras.*

Proof. (1) Given two crossed product algebras $C_0(X_i) \rtimes \Gamma_i$, a morphism $C_0(X_1) \rtimes \Gamma_1 \to C_0(X_2) \rtimes \Gamma_2$ is given by a proper continuous map $\pi : X_2 \to X_1$ and a proper group homomorphism $\alpha : \Gamma_1 \to \Gamma_2$, with the property that

$$\pi \circ \alpha(g) = g \circ \pi \quad \forall g \in \Gamma_1. \tag{3.434}$$

This determines an algebra homomorphism

$$C_0(X_1) \rtimes \Gamma_1 \ni \sum f_g U_g \mapsto \sum (f_g \circ \pi) U_{\alpha(g)} \in C_0(X_2) \rtimes \Gamma_2. \tag{3.435}$$

If the map π fails to be proper, the map of (3.435) only defines a homomorphism to the multiplier algebra

$$C_0(X_1) \rtimes \Gamma_1 \to M(C_0(X_2) \rtimes \Gamma_2). \tag{3.436}$$

In the case we are considering, the map π given by the determinant satisfies (3.434) since $\det \alpha(r) = r$, is continuous, but it is not proper, hence it gives rise to a morphism Υ to the multiplier algebra $M(C_0(M_2(\mathbb{A}_f) \times \mathrm{GL}_2(\mathbb{R})))$ as in (3.436) and (3.435).

(2) The extension $\tilde{\Upsilon}$ of Υ to multipliers gives a morphism

$$\tilde{\Upsilon} : C(\mathbb{A}_f \times \mathrm{GL}_1(\mathbb{R})) \rtimes \mathrm{GL}_1(\mathbb{Q}) \to C(M_2(\mathbb{A}_f) \times \mathrm{GL}_2(\mathbb{R})) \rtimes \mathrm{GL}_2(\mathbb{Q}).$$

One obtains Υ_0 by restricting $\tilde{\Upsilon}$ to homogeneous functions of weight 0 for the scaling action of $\mathbb{R}_+^*$ on $\mathrm{GL}_1(\mathbb{R})$. The image of such a function is homogeneous of weight 0 for the scaling action of $\mathbb{C}^*$ on $\mathrm{GL}_2(\mathbb{R})$ and only depends upon the sign of the modulus in $\mathbb{H}^{\pm}$ as in the diagram

$$\begin{tikzcd} \mathrm{GL}_2(\mathbb{R}) \arrow[r] \arrow[d, "\det"] & \mathbb{H}^{\pm} = \mathrm{GL}_2(\mathbb{R})/\mathbb{C}^* \arrow[d, "\mathrm{sign}"] \\ \mathrm{GL}_1(\mathbb{R}) \arrow[r] & \{\pm 1\} = \mathrm{GL}_1(\mathbb{R})/\mathbb{R}_+^*. \end{tikzcd} \tag{3.437}$$

(3) One can restrict Υ_0 to $\mathcal{A}_1 = C(\hat{\mathbb{Z}}) \rtimes \mathbb{N}$ and one gets a morphism

$$C(\hat{\mathbb{Z}}) \rtimes \mathbb{N} \to M(C_0(M_2(\mathbb{A}_f) \times \mathbb{H})) \rtimes \mathrm{GL}_2^+(\mathbb{Q}). \tag{3.438}$$

One then combines (3.438) with the Morita equivalence between $C_0(M_2(\mathbb{A}_f) \times \mathbb{H}) \rtimes \mathrm{GL}_2^+(\mathbb{Q})$ and $\mathcal{A}_2$. This Morita equivalence is given by the bimodule $\mathcal{E}$ of functions of

$$(g, \rho, z) \in \mathcal{V} = \mathrm{GL}_2^+(\mathbb{Q}) \times M_2(\hat{\mathbb{Z}}) \times \mathbb{H} \tag{3.439}$$

invariant under the Γ action

$$(g, \rho, z) \mapsto (g\gamma^{-1}, \gamma\rho, \gamma(z)). \tag{3.440}$$

To obtain the right module structure over $\mathcal{A}_2$, one uses the convolution product

$$(\xi * f)\,(g, \rho, z) = \sum_{s \in \Gamma\backslash \mathrm{GL}_2^+(\mathbb{Q})} \xi(gs^{-1}, s\,\rho, s(z))\, f(s, \rho, z). \tag{3.441}$$

For $(g, \rho, z) \in \mathcal{V}$, both $g\rho \in M_2(\mathbb{A}_f)$ and $g(z) \in \mathbb{H}$ make sense, and this can be endowed with a left action of $C_0(M_2(\mathbb{A}_f) \times \mathbb{H})$ on $\mathcal{E}$ given by

$$(f * \xi)\,(g, \rho, z) = f(g\rho, g(z))\, \xi(g, \rho, z)\,. \tag{3.442}$$

This action is $\mathrm{GL}_2^+(\mathbb{Q})$-equivariant for the left action of $\mathrm{GL}_2^+(\mathbb{Q})$ on $\mathcal{V}$ given by

$$h\,(g, \rho, z) = (h\,g, \rho, z) \quad \forall h \in \mathrm{GL}_2^+(\mathbb{Q}) \tag{3.443}$$

and turns $\mathcal{E}$ into a left $C_0(M_2(\mathbb{A}_f) \times \mathbb{H}) \rtimes \mathrm{GL}_2^+(\mathbb{Q})$ module.
The left action of $C_0(M_2(\mathbb{A}_f) \times \mathbb{H})$ extends to its multipliers and one can use (3.438) to obtain on $\mathcal{E}$ the required structure of left $C(\hat{\mathbb{Z}}) \rtimes \mathbb{N}$-module and right $\mathcal{A}_2$-module. One checks that this is compatible with the time evolutions σ_t. It is also compatible with the rational subalgebras $\mathcal{A}_{1,\mathbb{Q}}$ and $\mathcal{A}_{2,\mathbb{Q}}$. This compatibility can be seen using the cyclotomy condition of §7.2. $\square$

8. KMS states and complex multiplication

In this section we present a construction of a quantum statistical mechanical system that generalizes the Bost–Connes system in the case where instead of the field $\mathbb{Q}$ one considers an imaginary quadratic field $\mathbb{K}$. The main purpose is to recast the explicit class field theory for imaginary quadratic fields in the context of quantum statistical mechanics in much the same way as the Bost–Connes system does for the explicit class field theory of $\mathbb{Q}$. As we are going to see in the following, there is a very interesting interplay between the GL_2-system $(\mathcal{A}_2, \sigma_t)$ discussed in the previous section and the systems for imaginary quadratic fields, which mirrors the distinction between general elliptic curves and elliptic curves with complex multiplication. For this section we follow very closely the two papers of Connes–Marcolli–Ramachandran [**90**] and [**91**].

We write $\mathcal{O}$ for the ring of algebraic integers of the imaginary quadratic field $\mathbb{K} = \mathbb{Q}(\sqrt{-d})$, with d a positive integer. We can write the ring of integers in the form $\mathcal{O} = \mathbb{Z} + \mathbb{Z}\tau$, and the imaginary quadratic field as $\mathbb{K} = \mathbb{Q}(\tau)$, where we fix the embedding $\mathbb{K} \hookrightarrow \mathbb{C}$ so that τ is in the upper half plane $\mathbb{H}$. As in §4.5, we denote by $\mathbb{A}_{\mathbb{K}}$ the adèles of $\mathbb{K}$ and by $\mathbb{A}_{\mathbb{K},f}$ the finite adèles $\mathbb{A}_{\mathbb{K},f} = \mathbb{A}_{\mathbb{Q},f} \otimes_{\mathbb{Q}} \mathbb{K}$. We have $\mathbb{A}_{\mathbb{K}} = \mathbb{A}_{\mathbb{K},f} \times \mathbb{C}$. We denote here the maximal compact subring of $\mathbb{A}_{\mathbb{K},f}$ by $\hat{\mathcal{O}}$ instead of R as in §4.5. This is given by $\hat{\mathcal{O}} = \mathcal{O} \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}$. Moreover, we use the notation $\mathbb{A}^{\cdot}_{\mathbb{K}} = \mathbb{A}_{\mathbb{K},f} \times \mathbb{C}^*$.

8.1. 1-dimensional $\mathbb{K}$-lattices.

We introduce a generalization of the notion of $\mathbb{Q}$-lattices that we worked with in the previous sections, by considering an analogous definition of $\mathbb{K}$-lattices.

Definition 3.109. *Let $\mathbb{K}$ be an imaginary quadratic field. A 1-dimensional $\mathbb{K}$-lattice (Λ, ϕ) is a finitely generated $\mathcal{O}$-submodule $\Lambda \subset \mathbb{C}$ such that $\Lambda \otimes_{\mathcal{O}} \mathbb{K} \cong \mathbb{K}$, together with a morphism of $\mathcal{O}$-modules*

$$\phi : \mathbb{K}/\mathcal{O} \to \mathbb{K}\Lambda/\Lambda. \tag{3.444}$$

A 1-dimensional $\mathbb{K}$-lattice is invertible *if ϕ is an isomorphism of $\mathcal{O}$-modules.*
Two 1-dimensional $\mathbb{K}$-lattices (Λ_1, ϕ_1) and (Λ_2, ϕ_2) are commensurable if $\mathbb{K}\Lambda_1 = \mathbb{K}\Lambda_2$ and $\phi_1 = \phi_2$ modulo $\Lambda_1 + \Lambda_2$.

The proof that commensurability is an equivalence relation follows as in Lemma 3.18.
Here we can regard $\mathbb{C}$ as a $\mathbb{K}$-vector space and in particular as an $\mathcal{O}$-module and we assume that the $\mathcal{O}$-module structure of Λ is compatible with the embeddings of both $\mathcal{O}$ and Λ in $\mathbb{C}$.

Lemma 3.110. *As an $\mathcal{O}$-module, Λ is projective.*

Proof. As an $\mathcal{O}$-module Λ is isomorphic to a finitely generated $\mathcal{O}$-submodule of $\mathbb{K}$, hence to an ideal in $\mathcal{O}$. Every ideal in a Dedekind domain $\mathcal{O}$ is finitely generated projective over $\mathcal{O}$ (cf. [**230**] Lemma 1.5). ☐

Lemma 3.111. *Up to scaling by some $\lambda \in \mathbb{C}^*$, any $\mathbb{K}$-lattice Λ is equivalent to a $\mathbb{K}$-lattice $\Lambda' = \lambda\Lambda \subset \mathbb{K} \subset \mathbb{C}$. The lattice Λ' is unique modulo $\mathbb{K}^*$.*

Proof. The $\mathbb{K}$-vector space $\mathbb{K}\Lambda$ is 1-dimensional. If ξ is a generator, then $\xi^{-1}\Lambda \subset \mathbb{K}$. The remaining ambiguity is only by scaling by elements in $\mathbb{K}^*$. ☐

PROPOSITION 3.112. *For invertible 1-dimensional $\mathbb{K}$-lattices, the element of $K_0(\mathcal{O})$ associated to the $\mathcal{O}$-module Λ is an invariant of the commensurability class up to scaling.*

PROOF. Two invertible 1-dimensional $\mathbb{K}$-lattices that are commensurable are in fact equal. This is a special case of a stronger result proved below in Lemma 3.121 (1). The scaling of $\mathbb{K}$-lattices does not alter the isomorphism class of the $\mathcal{O}$-module Λ. Thus, the corresponding $\mathcal{O}$-module class is well defined. $\square$

There is a canonical isomorphism $K_0(\mathcal{O}) \cong \mathbb{Z} + \mathrm{Cl}(\mathcal{O})$ (cf. Corollary 1.11, [**229**]), where the $\mathbb{Z}$ part is given by the rank, which is equal to one in our case, hence the invariant of Proposition 3.112 is the class in the class group $\mathrm{Cl}(\mathcal{O})$. In contrast to Proposition 3.112, every 1-dimensional $\mathbb{K}$-lattice is commensurable to a $\mathbb{K}$-lattice whose $\mathcal{O}$-module structure is principal. This follows, since every ideal in $\mathcal{O}$ is commensurable to $\mathcal{O}$.

8.2. $\mathbb{K}$-lattices and $\mathbb{Q}$-lattices.

We can write the ring of integers of the imaginary quadratic field $\mathbb{K}$ in the form $\mathcal{O} = \mathbb{Z} + \mathbb{Z}\tau$, where τ is a point in the upper half plane $\mathbb{H}$. The field can then be written as $\mathbb{K} = \mathbb{Q}(\tau)$. The choice of such a generator $\tau \in \mathbb{H}$ determines (cf. [**268**] Proposition 4.6) an injective ring homomorphism

$$q_\tau : \mathbb{K} \hookrightarrow M_2(\mathbb{Q}). \tag{3.445}$$

In fact one has more precisely:

PROPOSITION 3.113. *For $\mathcal{O} = \mathbb{Z} + \mathbb{Z}\tau$ as above, let $q_\tau(k)$ be the matrix of multiplication by $k \in \mathbb{K}$ acting on the right on row vectors in the basis $\{\tau, 1\}$. Then $q_\tau : \mathbb{K} \hookrightarrow M_2(\mathbb{Q})$ is an injective ring homomorphism (embedding) such that*

$$q_\tau(\mathbb{K}^*) = \{g \in \mathrm{GL}_2^+(\mathbb{Q}) : \ g(\tau) = \tau\}\,. \tag{3.446}$$

There exist only two embeddings $\mathbb{K} \hookrightarrow M_2(\mathbb{Q})$ fulfilling (3.446). The second one q'_τ corresponds to the matrix of multiplication by $k \in \mathbb{K}$ acting on the left on column vectors in the basis $\{1, -\tau\}$.
They are related by

$$q'_\tau(k) = s\, q_\tau(k)^t\, s^{-1} = \, q_\tau(\bar{k})\,, \quad \forall k \in \mathbb{K}\,, \tag{3.447}$$

where $s = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $k \to \bar{k}$ is the automorphism of $\mathbb{K}$ given by complex conjugation.

PROOF. Since τ is a quadratic algebraic integer, it satisfies an equation of the form

$$\tau^2 - \alpha\tau + \beta = 0\,, \tag{3.448}$$

where $\alpha, \beta \in \mathbb{Z}$. One has by construction

$$q_\tau(x + \tau y) = \begin{pmatrix} x + \alpha y & -\beta y \\ y & x \end{pmatrix}\,, \quad \forall x\,,\, y \in \mathbb{Q}\,,$$

and one checks that provided $x + \tau y \neq 0$ one has $g = q_\tau(x + \tau y) \in \mathrm{GL}_2^+(\mathbb{Q})$ and $g(\tau) = \tau$. One has by construction

$$q'_\tau(x + \tau y) = \begin{pmatrix} x & \beta y \\ -y & x + \alpha y \end{pmatrix}, \quad \forall x\,,\ y \in \mathbb{Q}\,.$$

One checks similarly that provided $x + \tau y \neq 0$ one has $g' = q'_\tau(x + \tau y) \in \mathrm{GL}_2^+(\mathbb{Q})$ and $g'(\tau) = \tau$. Moreover the equality $q'_\tau(k) = s\, q_\tau(k)^t\, s^{-1}$ follows from a direct computation and just reflects the change of basis, and use of column vectors versus row vectors.
Finally any embedding $q : \mathbb{K} \hookrightarrow M_2(\mathbb{Q})$ satisfying (3.446) has range $q_\tau(\mathbb{K})$ and is therefore of the form $q = q_\tau \circ \sigma$ for an automorphism $\sigma \in \mathrm{Aut}(\mathbb{K})$, which gives the last equality of (3.447) and the required uniqueness. ☐

For $g = q_\tau(x)$ with $x \in \mathbb{K}^*$, we have

$$\det(g) = \mathfrak{n}(x), \tag{3.449}$$

where $\mathfrak{n} : \mathbb{K}^* \to \mathbb{Q}^*$ is the norm map.
The relation between 1-dimensional $\mathbb{K}$-lattices and 2-dimensional $\mathbb{Q}$-lattices is then explained by the following Lemma.

Lemma 3.114. *A 1-dimensional $\mathbb{K}$-lattice is, in particular, a 2-dimensional $\mathbb{Q}$-lattice. Two 1-dimensional $\mathbb{K}$-lattices are commensurable iff the underlying $\mathbb{Q}$-lattices are commensurable.*

Proof. First notice that $\mathbb{K}\Lambda = \mathbb{Q}\Lambda$, since $\mathbb{Q}\mathcal{O} = \mathbb{K}$. This, together with $\Lambda \otimes_{\mathcal{O}} \mathbb{K} \cong \mathbb{K}$, shows that the $\mathbb{Q}$-vector space $\mathbb{Q}\Lambda$ is 2-dimensional. Since $\mathbb{R}\Lambda = \mathbb{C}$, and Λ is finitely generated as an abelian group, this shows that Λ is a lattice. The map

$$(a, b) \in \mathbb{Q}^2/\mathbb{Z}^2 \to a - b\tau \in \mathbb{K}/\mathcal{O} \tag{3.450}$$

provides an identification of $\mathbb{Q}^2/\mathbb{Z}^2$ with $\mathbb{K}/\mathcal{O}$, so that we can view ϕ as a homomorphism of abelian groups $\phi : \mathbb{Q}^2/\mathbb{Z}^2 \to \mathbb{Q}\Lambda/\Lambda$. The pair (Λ, ϕ) thus gives a two-dimensional $\mathbb{Q}$-lattice.
The second statement holds, since for 1-dimensional $\mathbb{K}$-lattices we have $\mathbb{K}\Lambda = \mathbb{Q}\Lambda$. ☐

8.3. Adelic description of $\mathbb{K}$-lattices.

As we have seen in the case of $\mathbb{Q}$-lattices the noncommutative spaces that appear as algebras of observables of the corresponding quantum statistical mechanical systems are obtained as adelic quotients. This principle was systematically exploited in [**159**] to produce a large class of such systems associated to classical adelic quotients, namely Shimura varieties. In the case we are now considering of imaginary quadratic fields, it is also very useful to reformulate the notions of $\mathbb{K}$-lattices and commensurability in adelic terms.

Proposition 3.115. *The data (Λ, ϕ) of a 1-dimensional $\mathbb{K}$-lattice are equivalent to data (ρ, s) of an element $\rho \in \hat{\mathcal{O}}$ and $s \in \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$, modulo the $\hat{\mathcal{O}}^*$-action given by $(\rho, s) \mapsto (x^{-1}\rho, xs)$, $x \in \hat{\mathcal{O}}^*$. Thus, the space of 1-dimensional $\mathbb{K}$-lattices is given by*

$$\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*). \tag{3.451}$$

PROOF. Up to scaling, the $\mathcal{O}$-module Λ can be identified with an ideal in $\mathcal{O}$. These can be written in the form $s_f\hat{\mathcal{O}} \cap \mathbb{K}$ where $s_f \in \hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$ (cf. [**268**] §5.2). Thus, given a $\mathbb{K}$-lattice (Λ, ϕ), Λ can be described in the form $\Lambda_s = s_\infty^{-1}(s_f\hat{\mathcal{O}} \cap \mathbb{K})$, where $s = (s_f, s_\infty) \in \mathbb{A}^*_{\mathbb{K}}$. This satisfies $\Lambda_{ks} = \Lambda_s$ for all $k \in (\hat{\mathcal{O}}^* \times 1)\,\mathbb{K}^* \subset \mathbb{A}^*_{\mathbb{K}}$. If $\Lambda_s = \Lambda_{s'}$, then $s'_\infty s_\infty^{-1} \in \mathbb{K}^*$ and one can assume that $s'_\infty = s_\infty$. Thus one is reduced to the condition $s_f\hat{\mathcal{O}} \cap \mathbb{K} = s'_f\hat{\mathcal{O}} \cap \mathbb{K}$, which implies $s'_f s_f^{-1} \in \hat{\mathcal{O}}^*$. The data ϕ of the 1-dimensional $\mathbb{K}$-lattice can be described by the composite map $\phi = s_\infty^{-1}(s_f \circ \rho)$

$$
\begin{tikzcd}
\mathbb{K}/\mathcal{O} \arrow[r, "\rho"] \arrow[d, "\cong"] & \mathbb{K}/\mathcal{O} \arrow[r] \arrow[d, "\cong"] & \mathbb{K}/(s_f\hat{\mathcal{O}} \cap \mathbb{K}) \arrow[d, "\cong"] \\
\mathbb{A}_{\mathbb{K},f}/\hat{\mathcal{O}} \arrow[r, "\rho"] & \mathbb{A}_{\mathbb{K},f}/\hat{\mathcal{O}} \arrow[r, "s_f"] & \mathbb{A}_{\mathbb{K},f}/s_f\hat{\mathcal{O}}.
\end{tikzcd} \tag{3.452}
$$

where ρ is an element in $\hat{\mathcal{O}}$. By construction the map $(\rho, s) \mapsto (\Lambda_s, s_\infty^{-1}(s_f \circ \rho))$ passes to the quotient $\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*)$ and the above shows that it gives a bijection with the space of 1-dimensional $\mathbb{K}$-lattices. ☐

Notice that, even though Λ and $\mathcal{O}$ are not isomorphic as $\mathcal{O}$-modules, on the quotients we have $\mathbb{K}/\Lambda \simeq \mathbb{K}/\mathcal{O}$ as $\mathcal{O}$-modules, with the isomorphism realized by s_f in the diagram (3.452).

PROPOSITION 3.116. *Let $\mathbb{A}^{\cdot}_{\mathbb{K}} = \mathbb{A}_{\mathbb{K},f} \times \mathbb{C}^*$ be the subset of adèles of $\mathbb{K}$ with non-zero Archimedean component. The map $\Theta(\rho, s) = \rho\, s$,*

$$\Theta : \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*) \to \mathbb{A}^{\cdot}_{\mathbb{K}}/\mathbb{K}^*, \tag{3.453}$$

preserves commensurability and induces an identification of the set of commensurability classes of 1-dimensional $\mathbb{K}$-lattices (not up to scale) with the space $\mathbb{A}^{\cdot}_{\mathbb{K}}/\mathbb{K}^*$.

PROOF. The map is well defined, because ρs is invariant under the action $(\rho, s) \mapsto (x^{-1}\rho, xs)$ of $x \in \hat{\mathcal{O}}^*$. It is clearly surjective since $\mathbb{A}_{\mathbb{K},f} = \hat{\mathcal{O}}\,\mathbb{K}^*$. It remains to show that two $\mathbb{K}$-lattices have the same image if and only if they are in the same commensurability class. First we show that we can reduce to the case of principal K-lattices, without changing the value of the map Θ. Given a $\mathbb{K}$-lattice (Λ, ϕ), we write $\Lambda = \lambda J$, where $J \subset \mathcal{O}$ is an ideal, hence $\Lambda = \lambda(s_f\hat{\mathcal{O}} \cap \mathbb{K})$, where $\lambda = s_\infty^{-1} \in \mathbb{C}^*$ and $s_f \in \hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$. Then (Λ, ϕ) is commensurate to the principal $\mathbb{K}$-lattice $(\lambda\mathcal{O}, \phi)$. If (ρ, s) is the pair associated to (Λ, ϕ), with $s = (s_f, s_\infty)$ as above, then the corresponding pair (ρ', s') for $(\lambda\mathcal{O}, \phi)$ is given by $\rho' = s_f\,\rho$ and $s' = (1, s_\infty)$. Thus, we have $\Theta(\Lambda, \phi) = \Theta(\lambda\mathcal{O}, \phi)$. We can then reduce to proving the statement in the case of principal $\mathbb{K}$-lattices $(s_\infty^{-1}\mathcal{O}, s_\infty^{-1}\rho)$. In this case, the equality $(\rho, s_\infty) = k(\rho', s'_\infty)$, for $k \in \mathbb{K}^*$, means that we have $s_\infty = ks'_\infty$ and $\rho = k\rho'$. In turn, this is the relation of commensurability for principal $\mathbb{K}$-lattices. ☐

Thus, we obtain the following result for 1-dimensional $\mathbb{K}$-lattices.

LEMMA 3.117. *The map defined as $\Upsilon : (\Lambda, \phi) \mapsto \rho \in \hat{\mathcal{O}}/\mathbb{K}^*$ for principal $\mathbb{K}$-lattices extends to an identification, given by $\Upsilon : (\Lambda, \phi) \mapsto s_f\,\rho \in \mathbb{A}_{\mathbb{K},f}/\mathbb{K}^*$, of the set of* *commensurability classes of 1-dimensional K-lattices up to scaling with the quotient*

$$\hat{\mathcal{O}}/\mathbb{K}^* = \mathbb{A}_{\mathbb{K},f}/\mathbb{K}^*. \tag{3.454}$$

8.4. Algebra and time evolution.

We now describe the noncommutative algebra of coordinates of the space of commensurability classes of 1-dimensional $\mathbb{K}$-lattices up to scaling. We show that it is given by the convolution C^*-algebra of the groupoid of the equivalence relation.

We first consider the groupoid $\mathcal{G}_{1,\mathbb{K}}$ of the equivalence relation of commensurability on the space of 1-dimensional $\mathbb{K}$-lattices.
By construction, this groupoid is a subgroupoid of the groupoid $\mathcal{G}_2$ of commensurability classes of 2-dimensional $\mathbb{Q}$-lattices described in §5.2. Namely, we have the following result.

LEMMA 3.118. *The groupoid $\mathcal{G}_{1,\mathbb{K}}$ of the equivalence relation of commensurability of 1-dimensional $\mathbb{K}$-lattices is a subgroupoid of the groupoid $\mathcal{G}_2$ of commensurability of 2-dimensional $\mathbb{Q}$-lattices. Its structure as a locally compact étale groupoid is inherited from the embedding.*

PROOF. This follows directly from Lemma 3.114. $\square$

The groupoid $\mathcal{G}_{1,\mathbb{K}}$ corresponds to the quotient space $\mathbb{A}^{\cdot}_{\mathbb{K}}/\mathbb{K}^*$, hence its C^*-algebra is given, up to Morita equivalence, by the crossed product

$$C_0(\mathbb{A}^{\cdot}_{\mathbb{K}}) \rtimes \mathbb{K}^*. \tag{3.455}$$

When we consider 1-dimensional $\mathbb{K}$-lattices up to scaling and the commensurability relation on this set, the construction of the corresponding convolution algebra is more delicate. Passing to 1-dimensional $\mathbb{K}$-lattices up to scaling corresponds to taking the quotient $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ on the groupoid of the commensurability relation.

In the case of the GL_2-system, we encountered a problem in considering the analogous quotient $\mathcal{G}_2/\mathbb{C}^*$, because, as we have seen in Remark 3.41, the quotient $\mathcal{G}_2/\mathbb{C}^*$ is no longer a groupoid. This made it necessary to define the resulting convolution algebra not as a groupoid C^*-algebra for $\mathcal{G}_2/\mathbb{C}^*$ but as a subalgebra of homogeneous elements of degree zero with respect to the $\mathbb{C}^*$-action in $C^*(\mathcal{G}_2)$, as discussed in §5.2.

In the case of 1-dimensional $\mathbb{K}$-lattices, the situation is better, and closer to the case of 1-dimensional $\mathbb{Q}$-lattices. In fact, we have the following result.

PROPOSITION 3.119. *The quotient $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ is a groupoid.*

PROOF. The simplest way to check this is to write $\mathcal{G}_{1,\mathbb{K}}$ as the union of the two groupoids $\mathcal{G}_{1,\mathbb{K}} = \mathcal{G}_0 \cup \mathcal{G}_1$ corresponding respectively to pairs of commensurable $\mathbb{K}$-lattices (L, L') with $L = (\Lambda, 0), L' = (\Lambda', 0)$ and (L, L') with $L = (\Lambda, \phi \neq 0), L' = (\Lambda', \phi' \neq 0)$. For $\lambda \in \mathbb{C}^*$ one has $\lambda\Lambda = \Lambda$ iff $\lambda \in \mathcal{O}^*$ is an invertible element of $\mathcal{O}$. Thus the scaling action of $\lambda \in \mathbb{C}^*$ on $\mathcal{G}_0$ is the identity for λ in the finite group $\mathcal{O}^*$ of invertible elements of $\mathcal{O}$ and the corresponding action of $\mathbb{C}^*/\mathcal{O}^*$ is free on the units of $\mathcal{G}_0$. Thus the quotient $\mathcal{G}_0/\mathbb{C}^*$ is a groupoid. Similarly the action of $\mathbb{C}^*$ on $\mathcal{G}_1$ is free on the units of $\mathcal{G}_1$ since for $\lambda \in \mathcal{O}^*$, $\lambda \neq 1$ the element $\lambda - 1 \in \mathbb{K}$ is invertible, so that $\lambda\phi \neq \phi$. Thus the quotient $\mathcal{G}_1/\mathbb{C}^*$ is a groupoid. $\square$

The quotient topology turns $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ into a locally compact étale groupoid. The C^*-algebra of coordinates $\mathcal{A}_{1,\mathbb{K}} = C^*(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)$ is obtained from the convolution algebra of functions on $\mathcal{G}_{1,\mathbb{K}}$ that are of weight zero with respect to the $\mathbb{C}^*$-action. Namely, one obtains elements in the algebra $\mathcal{A}_{1,\mathbb{K}}$ from functions of pairs $(\Lambda,\phi),(\Lambda',\phi')$ of commensurable 1-dimensional $\mathbb{K}$-lattices satisfying

$$f(\lambda(\Lambda,\phi),\lambda(\Lambda',\phi')) = f((\Lambda,\phi),(\Lambda',\phi')) \quad \forall\lambda\in\mathbb{C}^*.$$

Unlike the convolution algebra $\mathcal{A}_2$ of the GL_2 system, the algebra $\mathcal{A}_{1,\mathbb{K}}$ is unital and resembles more closely the Bost–Connes algebra $\mathcal{A}_1$.

LEMMA 3.120. *The algebra $\mathcal{A}_{1,\mathbb{K}}$ is unital.*

PROOF. The set $(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)^{(0)}$ of units of the groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ is the quotient of $\hat{\mathcal{O}}\times_{\hat{\mathcal{O}}^*}(\mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*)$ by the action of $\mathbb{C}^*$. This gives the compact space

$$(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)^{(0)} = \hat{\mathcal{O}}\times_{\hat{\mathcal{O}}^*}(\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*). \tag{3.456}$$

Notice that $\mathbb{A}^*_{\mathbb{K},f}/(\mathbb{K}^*\times\hat{\mathcal{O}}^*)$ is the class group $\mathrm{Cl}(\mathcal{O})$. Since the set of units of the groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ is compact the convolution algebra is unital. □

There is a homomorphism $\mathfrak{n}$ from the groupoid $\mathcal{G}_{1,\mathbb{K}}$ to $\mathbb{R}^*_+$ given by the covolume of a commensurable pair of $\mathbb{K}$-lattices. More precisely, given such a pair $L=(\Lambda,\phi)\sim L'=(\Lambda',\phi')$, we let, as in (3.57),

$$\mathfrak{n}(L,L') = |L/L'| = \frac{\mathrm{covolume}(\Lambda')}{\mathrm{covolume}(\Lambda)}. \tag{3.457}$$

This coincides with the order of the quotient L/L' when $L'\subset L$. It is invariant under scaling, hence it descends to the quotient groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$.
This defines a time evolution on the algebra $\mathcal{A}_{1,\mathbb{K}}$ by

$$\sigma_t(f)(L,L') = |L/L'|^{it} f(L,L'). \tag{3.458}$$

Up to scale, we can identify the lattices in a commensurable pair with ideals in $\mathcal{O}$. For $L=(\Lambda,\phi)\sim L'=(\Lambda',\phi')$ as above with $\Lambda=\lambda I$ and $\Lambda'=\lambda J$, the covolume is given by the ratio of the norms of the ideals as

$$\mathfrak{n}(L,L') = \frac{\mathfrak{n}(J)}{\mathfrak{n}(I)}. \tag{3.459}$$

(recall that the norm $\mathfrak{n}(J)$ of an ideal J is the cardinality of $\mathcal{O}/J$, which gives (3.459) for $\Lambda = I = \mathcal{O}$ and $\Lambda' = J$ an ideal).

In order to construct representations of the algebra $\mathcal{A}_{1,\mathbb{K}}$, we adopt the same technique described in §4. Namely, we use the fact that, for an étale groupoid like our $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$, every unit $y\in(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)^{(0)}$ defines by (3.43) a representation π_y by left convolution of the algebra of $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ in the Hilbert space $\mathcal{H}_y = \ell^2((\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)_y)$, where $(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)_y$ denotes the set of elements of the groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ with source y. As in the GL_2 case, this defines the norm on the C^*-algebra $\mathcal{A}_{1,\mathbb{K}}$ as

$$\|f\| = \sup_{y\in(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)^{(0)}} \|\pi_y(f)\|. \tag{3.460}$$

8.5. $\mathbb{K}$-lattices and ideals.

The following result gives a description of the commensurability class of an *invertible* 1-dimensional $\mathbb{K}$-lattice in terms of ideals in $\mathcal{O}$.

LEMMA 3.121. (1) *Given an invertible 1-dimensional $\mathbb{K}$-lattice (Λ, ϕ), the map*

$$(\Lambda', \phi') \mapsto J = \{x \in \mathcal{O} | x\Lambda' \subset \Lambda\} \tag{3.461}$$

gives a bijection of the set of 1-dimensional $\mathbb{K}$-lattices up to scale commensurable to (Λ, ϕ) with the set of (nonzero) ideals in $\mathcal{O}$.

(2) *If L is an invertible $\mathbb{K}$-lattice, the corresponding representation π_L of the algebra $\mathcal{A}_{1,\mathbb{K}}$ has positive energy.*

(3) *The partition function $Z_{\mathbb{K}}(\beta)$ is the Dedekind zeta function of the imaginary quadratic field $\mathbb{K}$,*

$$Z_{\mathbb{K}}(\beta) = \zeta_{\mathbb{K}}(\beta). \tag{3.462}$$

PROOF. (1) We use the result of Lemma 3.93, by which we know that, if Λ is an invertible 2-dimensional $\mathbb{Q}$-lattice and Λ' is commensurable to Λ, then $\Lambda \subset \Lambda'$. Since 1-dimensional $\mathbb{K}$-lattices are also 2-dimensional $\mathbb{Q}$-lattices, the result applies. The map (3.461) is well defined, since $J \subset \mathcal{O}$ is an ideal. Moreover, $J\Lambda' = \Lambda$, since $\mathcal{O}$ is a Dedekind domain (one can assume that Λ and Λ' are ideals in $\mathcal{O}$ and use the definition of Dedekind domains of [**230**]). The map is injective, since J determines Λ' as the $\mathcal{O}$-module $\{x \in \mathbb{C} | xJ \subset \Lambda\}$ and the corresponding ϕ and ϕ' agree. This also shows that the map is surjective.
We introduce the notation

$$J^{-1}(\Lambda, \phi) = (\Lambda', \phi). \tag{3.463}$$

(2) For an invertible $\mathbb{K}$-lattice L, we have obtained in this way an identification of $(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)_L$ with the set $\mathcal{J}_{\mathcal{O}}$ of (nonzero) ideals $J \subset \mathcal{O}$. We let ϵ_J be the basis vector associated to $(J^{-1}L, L) \in (\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)_L$. The covolume of the pair $(J^{-1}L, L)$ is then given by the norm $\mathfrak{n}(J)$ of the ideal. The representation (3.43) is of the form

$$(\pi_L(f)\xi)(g) = \sum_{g_1 g_2 = g} f(g_1)\xi(g_2).$$

Thus the time evolution (3.458) is implemented by $\xi(g) \to \mathfrak{n}(g)^{it}\, \xi(g)$ and the corresponding Hamiltonian is of the form

$$H\epsilon_J = \log \mathfrak{n}(J)\ \epsilon_J, \tag{3.464}$$

with non-negative eigenvalues. Thus, the representation has positive energy.
(3) The partition function of the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ is then given by the Dedekind zeta function

$$Z(\beta) = \mathrm{Tr}(e^{-\beta H}) = \sum_{J \in \mathcal{J}_{\mathcal{O}}} \mathfrak{n}(J)^{-\beta} = \zeta_{\mathbb{K}}(\beta). \tag{3.465}$$

□

We give a more explicit description of the action of ideals $J \in \mathcal{J}_{\mathcal{O}}$ on $\mathbb{K}$-lattices described in the proof of Lemma 3.121 above.

LEMMA 3.122. *Let $L = (\Lambda, \phi)$ be a 1-dimensional $\mathbb{K}$-lattice and $J \subset \mathcal{O}$ an ideal. If L is represented by a pair $(\rho, s) \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} \mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*$, then the $\mathbb{K}$-lattice $J^{-1}L$ is represented by the commensurable pair $(s_J\rho, s_J^{-1}s)$, where s_J is a finite idèle such that $J = s_J\hat{\mathcal{O}} \cap \mathbb{K}$.*

PROOF. By [**268**] §5.2, there is a finite idèle s_J such that we can write the ideal J in the form $J = s_J\hat{\mathcal{O}} \cap \mathbb{K}$. The pair $(s_J\rho, s_J^{-1}s)$ defines an element in $\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} \mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*$. In fact, first notice that s_J is in fact in $\hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$, hence the product $s_J\rho \in \hat{\mathcal{O}}$. It is well defined modulo $\hat{\mathcal{O}}^*$, and by direct inspection one sees that the class it defines in $\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} \mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*$ is that of $J^{-1}L$. In fact with $s = (s_f, s_\infty)$ the lattice Λ_s is $s_\infty^{-1}(s_f\hat{\mathcal{O}} \cap \mathbb{K})$ and the lattice $\Lambda_{s_J^{-1}s}$ is $s_\infty^{-1}(s_J^{-1}s_f\hat{\mathcal{O}} \cap \mathbb{K})$ which agrees with $J^{-1}\Lambda_s$. Moreover $s_\infty^{-1}(s_f \circ \rho) = s_\infty^{-1}(s_f s_J^{-1} \circ s_J\rho)$. By Proposition 3.116, in order to check that the $\mathbb{K}$-lattices (ρ, s) and $(s_J\rho, s_J^{-1}s)$ lie in the same commensurability class, it is sufficient to see that $\Theta(\rho, s) = \rho s = \Theta(s_J\rho, s_J^{-1}s)$. $\square$

8.6. Arithmetic subalgebra.

The choice of an arithmetic subalgebra $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}}$ for the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ is obtained using the fact that the system can be embedded as a subsystem of the GL_2-system $(\mathcal{A}_2, \sigma_t)$, hence one can restrict the arithmetic algebra $\mathcal{A}_{2,\mathbb{Q}}$ of the GL_2-system to the subgroupoid $\mathcal{G}_{1,\mathbb{K}} \subset \mathcal{G}_2$ of commensurability classes of 1-dimensional $\mathbb{K}$-lattices. In fact, we have the following result.

LEMMA 3.123. *The time evolution* (3.458) *on the C^*-algebra $\mathcal{A}_{1,\mathbb{K}}$ is the restriction of the time evolution on $\mathcal{A}_2$ to the subgroupoid $\mathcal{G}_{1,\mathbb{K}} \subset \mathcal{G}_2$.*

PROOF. This is clear from Lemma 3.114 and from Lemma 3.118 above, together with the description in (3.206) of the time evolution of the system $(\mathcal{A}_2, \sigma_t)$ in terms of covolume of 2-dimensional lattices, or equivalently the description (3.449) of the norm map. $\square$

Thus, we have shown that the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ is a subsystem of $(\mathcal{A}_2, \sigma_t)$ obtained by restricting the functions of $\mathcal{A}_2$ to $\mathcal{G}_{1,\mathbb{K}}$ and with the induced time evolution.

We denote by $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}}$ the algebra obtained by restricting elements of $\mathcal{A}_{2,\mathbb{Q}}$ to the groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$. The algebra $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}}$ obtained in this way is defined over $\mathbb{Q}$. To have later the correct Galois properties, we need to consider as arithmetic algebra for the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ the algebra over $\mathbb{K}$

$$\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{K}. \tag{3.466}$$

Since $\mathcal{A}_{1,\mathbb{K}}$ is unital, one sees easily that $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{C}$ is a *subalgebra* of $\mathcal{A}_{1,\mathbb{K}}$, even though the arithmetic algebra $\mathcal{A}_{2,\mathbb{Q}}$ is only an algebra of unbounded multipliers of $\mathcal{A}_2$.

The relations described in §7.3 that hold in the arithmetic algebra $\mathcal{A}_{2,\mathbb{Q}}$ induce corresponding relations in the algebra $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$. Even though the relations described in §7.3 did not give a complete set of relations for $\mathcal{A}_{2,\mathbb{Q}}$, these are sufficient to show that the abelian part of the algebra $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$ yields an interesting AF algebra, like in the Bost–Connes system. This will become relevant in the following chapter in the context of endomotives.

8.7. Symmetries.

Let $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ be the system introduced above, for $\mathbb{K}$ an imaginary quadratic field. We have the following symmetries ([**90**]).

PROPOSITION 3.124. *The semigroup $\hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$ acts on the algebra $\mathcal{A}_{1,\mathbb{K}}$ by endomorphisms. The subgroup $\hat{\mathcal{O}}^*$ acts on $\mathcal{A}_{1,\mathbb{K}}$ by automorphisms. The subsemigroup $\mathcal{O}^\times$ acts by inner endomorphisms.*

PROOF. Given an ideal $J \subset \mathcal{O}$, consider the set of $\mathbb{K}$-lattices (Λ, ϕ) such that ϕ is well defined modulo $J\Lambda$. Namely, the map $\phi : \mathbb{K}/\mathcal{O} \to \mathbb{K}\Lambda/\Lambda$ factorises through $\mathbb{K}/\mathcal{O} \to \mathbb{K}\Lambda/J\Lambda \to \mathbb{K}\Lambda/\Lambda$. We say, in this case, that the $\mathbb{K}$-lattice (Λ, ϕ) is divisible by J. This condition gives a closed and open subset E_J of the set of $\mathbb{K}$-lattices up to scaling. We denote by $e_J \in \mathcal{A}_{1,\mathbb{K}}$ the corresponding idempotent. Let $s \in \hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$. Let $J = s\hat{\mathcal{O}} \cap \mathbb{K}$. Then for $(\rho, t) \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*)$ (cf. Proposition 3.115) the associated 1-dimensional $\mathbb{K}$-lattice (Λ, ϕ) is divisible by J iff $s^{-1}\rho \in \hat{\mathcal{O}}$. We denote the $\mathbb{K}$-lattice corresponding to $(s^{-1}\rho, t) \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*)$ simply by $(\Lambda, s^{-1}\phi)$. This map preserves commensurability and defines an isomorphism of the étale groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ with its reduction by E_J.
Given a commensurable pair (Λ, ϕ) and (Λ', ϕ'), and an element $f \in \mathcal{A}_{1,\mathbb{K}}$, we define

$$\theta_s(f)((\Lambda, \phi), (\Lambda', \phi')) = f((\Lambda, s^{-1}\phi), (\Lambda', s^{-1}\phi')) \tag{3.467}$$

if both $\mathbb{K}$-lattices are divisible by J and

$$\theta_s(f)((\Lambda, \phi), (\Lambda', \phi')) = 0 \quad \text{otherwise.}$$

Formula (3.467) defines an endomorphism of $\mathcal{A}_{1,\mathbb{K}}$ with range the algebra reduced by e_J. It is, in fact, an isomorphism with the reduced algebra. Clearly, for $s \in \hat{\mathcal{O}}^*$, the above defines an automorphism. For $s \in \mathcal{O}^\times$, the endomorphism (3.467) is inner. In fact, for $s \in \mathcal{O}^\times$, let $\mu_s \in \mathcal{A}_{1,\mathbb{K}}$ be given by

$$\mu_s((\Lambda, \phi), (\Lambda', \phi')) = \begin{cases} 1 & \Lambda = s^{-1}\Lambda' \text{ and } \phi' = \phi \\ 0 & \text{otherwise.} \end{cases} \tag{3.468}$$

Then the range of μ_s is the projection e_J, where J is the principal ideal generated by s. Then we have

$$\theta_s(f) = \mu_s\, f\, \mu_s^*, \quad \forall s \in \mathcal{O}^\times.$$

The action of symmetries $\hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$ is compatible with the time evolution,

$$\theta_s\, \sigma_t = \sigma_t\, \theta_s, \qquad \forall s \in \hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f},\ \forall t \in \mathbb{R}.$$

The isometries μ_s are eigenvectors of the time evolution, namely

$$\sigma_t(\mu_s) = \mathfrak{n}(s)^{it}\, \mu_s.$$

□

The induced symmetries of KMS states of the system are then described by the following results.

PROPOSITION 3.125. *The quotient $C_{\mathbb{K}}/D_{\mathbb{K}}$ of the idèle class group $C_{\mathbb{K}}$ of $\mathbb{K}$ by the connected component of the identity $D_{\mathbb{K}}$ acts as symmetries of the KMS states of the system $(\mathcal{A}_{\mathbb{K}}, \sigma_t)$. The action of the subgroup $\hat{\mathcal{O}}^*/\mathcal{O}^*$ is by automorphisms.*

PROOF. Recall that $\mathbb{A}_{\mathbb{K},f} = \hat{\mathcal{O}}\mathbb{K}^*$. Thus, we can pass to the corresponding group of symmetries, modulo inner, which is given by the group $\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$, which is isomorphic to $C_{\mathbb{K}}/D_{\mathbb{K}}$ (cf. [**268**] §5, [**6**] §9). We have an exact sequence of groups

$$1 \to \hat{\mathcal{O}}^*/\mathcal{O}^* \to \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^* \to \mathrm{Cl}(\mathcal{O}) \to 1,$$

where $\mathrm{Cl}(\mathcal{O})$ is the class group of the ring $\mathcal{O}$, with $\#\mathrm{Cl}(\mathcal{O}) = h_{\mathbb{K}}$, the class number of $\mathbb{K}$. ☐

This shows that, in the very special case of class number $h_{\mathbb{K}} = 1$, symmetries are given only through an action by automorphisms, as in the original case of the Bost–Connes system. In the case where $h_{\mathbb{K}} \neq 1$, the nontrivial elements of $\mathrm{Cl}(\mathcal{O})$ have representatives in $\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ that act by endomorphisms.

Notice that we can consider isometries $\mu_J \in \mathcal{A}_{1,\mathbb{K}}$ with range e_J, the characteristic function of the set of $\mathbb{K}$-lattices divisible by J. This isometry is of the form

$$\mu_J((\Lambda,\phi),(\Lambda',\phi')) = \begin{cases} 1 & \Lambda = J^{-1}\Lambda' \text{ and } \phi' = \phi \\ 0 & \text{otherwise.} \end{cases} \tag{3.469}$$

This is similar to equation (3.468) and it reduces to (3.468) in the case where J is principal (generated by s). Thus, this would seem to imply that it is not only the subsemigroup $\mathcal{O}^\times$ that acts by inner endomorphisms, but in fact a bigger one, obtained by using the inner action of the isometries $\mu_J \in \mathcal{A}_{1,\mathbb{K}}$. To see why this is not the case, one needs to compare the endomorphism $f \mapsto \mu_J f \mu_J^*$ with the endomorphism θ_s. In the first case one gets

$$\mu_J f \mu_J^*((\Lambda,\phi),(\Lambda',\phi')) = f((J\,\Lambda,\phi),(J\,\Lambda',\phi')) \tag{3.470}$$

if both $\mathbb{K}$-lattices are divisible by J and

$$\mu_J f \mu_J^*((\Lambda,\phi),(\Lambda',\phi')) = 0 \quad \text{otherwise,}$$

while in the second case one gets the formula (3.467) i.e.

$$\theta_s(f)((\Lambda,\phi),(\Lambda',\phi')) = f((\Lambda,s^{-1}\phi),(\Lambda',s^{-1}\phi')) \tag{3.471}$$

if both $\mathbb{K}$-lattices are divisible by J and

$$\theta_s(f)((\Lambda,\phi),(\Lambda',\phi')) = 0 \quad \text{otherwise.}$$

The key point here is that the scaling is only allowed by elements of $\mathbb{K}^*$ and the scaling relation between lattices $(s\,\Lambda,\phi)$ and $(\Lambda,s^{-1}\phi)$ holds only for $s \in \mathbb{K}^*$, but not for more general idèles. Thus, even though the isometries μ_J always exist for any ideal J, their inner action implements the endomorphism θ_s only in the case where J is principal.

COROLLARY 3.126. *The symmetries of the system $(\mathcal{A}_{1,\mathbb{K}},\sigma_t)$ preserve the arithmetic algebra $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$.*

PROOF. This follows from Proposition 3.86. In fact, one can see that these symmetries are induced by symmetries of the system $(\mathcal{A}_2,\sigma_t)$, through the diagram with exact rows

$$\begin{tikzcd} 1 \arrow[r] & \mathbb{K}^* \arrow[r] & \mathrm{GL}_1(\mathbb{A}_{\mathbb{K},f}) \arrow[r] \arrow[d, "q'_\tau"] & C_{\mathbb{K}}/D_{\mathbb{K}} \arrow[r] & 1 \\ 1 \arrow[r] & \mathbb{Q}^* \arrow[r] & \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \arrow[r] & S = \mathbb{Q}^*\backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \arrow[r] & 1. \end{tikzcd} \tag{3.472}$$

More precisely, the symmetry θ_s of the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$, associated to $s \in \mathrm{GL}_1(\mathbb{A}_{\mathbb{K},f})$ is the restriction of the symmetry ϑ_g (cf. (3.52)) of the system $(\mathcal{A}_2, \sigma_t)$ for

$$g = q'_\tau(s^{-1}) \in \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \,. \tag{3.473}$$

This follows, comparing the definition (3.467) of θ_s with (3.52) and using the defining property (cf. Proposition 3.113) of $q'_\tau(s^{-1})$ as the matrix of multiplication by s^{-1} in the basis $\{1, -\tau\}$. $\square$

8.8. Low-temperature KMS states and Galois action.

The main result on low-temperature KMS states of the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ is the following ([**90**]).

THEOREM 3.127. *The extremal KMS states of this system satisfy:*

- *For $\beta > 1$, extremal KMS$_\beta$ states are parameterized by the invertible $\mathbb{K}$-lattices,*

$$\mathcal{E}_\beta \simeq \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^* \tag{3.474}$$

  with a free and transitive action of $C_\mathbb{K}/D_\mathbb{K} \cong \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ as symmetries.
- *In this range, the extremal KMS$_\beta$ state associated to an invertible $\mathbb{K}$-lattice $L = (\Lambda, \phi)$ is of the form*

$$\varphi_{\beta,L}(f) = \zeta_\mathbb{K}(\beta)^{-1} \sum_{J \in \mathcal{J}_\mathcal{O}} f(J^{-1}L, J^{-1}L)\; \mathfrak{n}(J)^{-\beta}, \tag{3.475}$$

  where $\zeta_\mathbb{K}(\beta)$ is the Dedekind zeta function, and $J^{-1}L$ defined as in (3.463)*.*
- *The set of extremal KMS$_\infty$ states (as weak limits of KMS$_\beta$ states) is still given by* (3.474)*.*
- *The extremal KMS$_\infty$ states $\varphi_{\infty,L}$ of the CM system, evaluated on the arithmetic subalgebra $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$, take values in $\mathbb{K}^{\mathrm{ab}}$, with*

$$\varphi_{\infty,L}(\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}) = \mathbb{K}^{\mathrm{ab}}.$$

- *The class field theory isomorphism* (3.127) *intertwines the action of $\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ by symmetries of the system $(\mathcal{A}_\mathbb{K}, \sigma_t)$ and the action of $\mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K})$ on the image of $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$ under the extremal KMS$_\infty$ states. Namely, for all $\varphi_{\infty,L} \in \mathcal{E}_\infty$ and for all $f \in \mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$,*

$$\alpha(\varphi_{\infty,L}(f)) = (\varphi_{\infty,L} \circ \theta^{-1}(\alpha))(f), \qquad \forall \alpha \in \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K}). \tag{3.476}$$

PROOF. The partition function $Z_\mathbb{K}(\beta)$ of (3.465) converges for $\beta > 1$. We have also seen in the previous section that invertible $\mathbb{K}$-lattices $L = (\Lambda, \phi)$ determine positive energy representations of $\mathcal{A}_{1,\mathbb{K}}$ on the Hilbert space $\mathcal{H} = \ell^2(\mathcal{J})$ where $\mathcal{J}$ is the set of ideals of $\mathcal{O}$. Thus, the formula

$$\varphi_{\beta,L}(f) = \frac{\mathrm{Tr}\,(\pi_L(f)\,\exp(-\beta H))}{\mathrm{Tr}(\exp(-\beta H))} \tag{3.477}$$

defines an extremal KMS$_\beta$ state, with the Hamiltonian H of (3.464). These states are of the form (3.475). It is not hard to see that distinct elements in $\mathbb{A}_{\mathbb{K},f}/\mathbb{K}^*$ define distinct states $\varphi_{\beta,L}$.

This shows that we have an injection of $\mathbb{A}_{\mathbb{K},f}/\mathbb{K}^* \subset \mathcal{E}_\beta$. We need to show that, conversely, every extremal KMS$_\beta$ state is of the form (3.475).

We proceed in two steps. The first shows (Proposition 3.130 below) that KMS_β states are given by measures on the space of $\mathbb{K}$-lattices up to scaling, as in (3.478) below. The second shows that, when $\beta > 1$, this measure is carried by the commensurability classes of invertible $\mathbb{K}$-lattices.

Let p denotes the projection from $\mathbb{K}$-lattices to their class $p(L)$ modulo scaling, namely

$$p : \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*) \to \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*).$$

We obtain the following result.

LEMMA 3.128. *Let $\gamma = (L, L') \in \mathcal{G}_{1,\mathbb{K}}$ with*

$$p(L) = p(L') \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*).$$

Then either $L = L'$ or $\phi = \phi' = 0$.

PROOF. We first describe the elements $\gamma \in \mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ such that $s(\gamma) = r(\gamma)$, i.e. $\gamma = (L, L') \in \mathcal{G}_{1,\mathbb{K}}$ such that the classes of L and L' modulo the scaling action of $\mathbb{C}^*$ are the same elements of the set $(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)^{(0)}$ of units as in (3.456). Modulo scaling, we can assume that the lattice $\Lambda \subset \mathbb{K}$. Thus, since L and L' are commensurable, it follows that $\Lambda' \subset \mathbb{K}$. Then, by hypothesis, there exists $\lambda \in \mathbb{C}^*$ such that $\lambda L = L'$. One has $\lambda \in \mathbb{K}^*$ and $\phi' = \lambda\phi$. By the commensurability of the pair one also has $\phi' = \phi$ modulo $\Lambda + \Lambda'$. Writing $\lambda = \frac{a}{b}$, with $a, b \in \mathcal{O}$, we get $a\phi = b\phi$ and $\lambda = 1$ unless $\phi = 0$. $\square$

We let Ξ be the finite closed subset of $\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$ given by the set of $\mathbb{K}$-lattices up to scaling such that $\phi = 0$. Since $\mathbb{A}^*_{\mathbb{K},f}/\hat{\mathcal{O}}^*\mathbb{K}^* = \mathrm{Cl}(\mathcal{O})$, its cardinality $\#\Xi = h_{\mathbb{K}}$ is the class number of $\mathbb{K}$. The groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ is the union $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^* = \mathcal{G}_0 \cup \mathcal{G}_1$ of the reduced groupoids by $\Xi \subset \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$ and its complement.

LEMMA 3.129. *Let $\gamma \in \mathcal{G}_1 \backslash \mathcal{G}_1^{(0)}$. There exists a neighborhood V of γ in $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ such that*

$$r(V) \cap s(V) = \emptyset$$

where r and s are the range and source maps of $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$.

PROOF. Let γ be the class modulo scaling of the commensurable pair (L, L'). By Lemma 3.128 one has $p(L) \neq p(L') \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$. Let $c(\Lambda)$ and $c(\Lambda')$ denote the classes of Λ and Λ' in $K_0(\mathcal{O})$ (cf. Proposition 3.112). If these classes are different, $c(\Lambda) \neq c(\Lambda')$, then one can simply take a neighborhood V in such a way that all elements $\gamma_1 = (L_1, L_1') \in V$ are in the corresponding classes: $c(\Lambda_1) = c(\Lambda)$, $c(\Lambda_1') = c(\Lambda')$, for $L_1 = (\Lambda_1, \phi_1)$ and $L_1' = (\Lambda_1', \phi_1')$. This ensures that range and source are disjoint sets, $r(V) \cap s(V) = \emptyset$. Otherwise, there exists $\lambda \in \mathbb{K}^*$ such that $\Lambda' = \lambda\Lambda$. Since $\Lambda' \neq \Lambda$, one has $\lambda \notin \mathcal{O}^*$. One has $\phi' = \phi \neq 0$. Thus, one is reduced to showing that, given $\rho \in \hat{\mathcal{O}}$, $\rho \neq 0$, and $\lambda \in \mathbb{K}^*$, $\lambda \notin \mathcal{O}^*$, there exists a neighborhood W of ρ in $\hat{\mathcal{O}}$ such that $\lambda W \cap \mathcal{O}^* W = \emptyset$. This follows, using a place v such that $\rho_v \neq 0$. In fact, one has $\lambda\rho_v \notin \mathcal{O}^*\rho_v$ and the same holds in a suitable neighborhood. $\square$

We can now prove the following.

PROPOSITION 3.130. *Let $\beta > 0$ and let φ be a KMS$_\beta$ state for $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$. Then there exists a probability measure μ on $\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$ such that*

$$\varphi(f) = \int_{\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)} f(L,L)\, d\mu(L), \quad \forall f \in \mathcal{A}_{1,\mathbb{K}}\,. \tag{3.478}$$

PROOF. It is enough to show that $\varphi(f) = 0$ provided f is a continuous function with compact support on $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$ with support disjoint from $(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)^{(0)}$. Let $h_n \in C(\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*))$ with $0 \leq h_n \leq 1$ and with support disjoint from Ξ and converging pointwise to 1 in the complement of Ξ. Let $u_n \in \mathcal{A}_{1,\mathbb{K}}$ be supported by the diagonal and given by h_n there.
The formula

$$\Phi(f)(\Lambda, \Lambda') := f((\Lambda, 0), (\Lambda', 0))) \quad \forall f \in \mathcal{A}_{1,\mathbb{K}} \tag{3.479}$$

defines a homomorphism of the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ to the C^* dynamical system $(C^*(\mathcal{G}_0), \sigma_t)$ obtained by specialization to pairs of $\mathbb{K}$-lattices with $\phi = 0$ as in Theorem 3.95.
Since there are unitary eigenvectors for σ_t for nontrivial eigenvalues in the system $(C^*(\mathcal{G}_0), \sigma_t)$ it has no nonzero KMS$_\beta$ positive functional. This shows that the pushforward of φ by Φ vanishes and as in Theorem 3.95 that, with the notation introduced above,

$$\varphi(f) = \lim_n \varphi(f * u_n).$$

Thus, since $(f * u_n)(\gamma) = f(\gamma)\, h_n(s(\gamma))$, we can assume that $f(\gamma) = 0$ unless $s(\gamma) \in \Omega$, where $\Omega \subset \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$ is a compact subset disjoint from Ξ. Let $L \in \Omega$ and V as in Lemma 3.129 and let $h \in C_c(V)$. Then, we apply the KMS$_\beta$ condition to the pair a, b, with $a = f$ and b supported by the diagonal and equal to h there. We get $\varphi(b * f) = \varphi(f * b)$. One has $(b * f)(\gamma) = h(r(\gamma))\, f(\gamma)$. Applying this to $f * b$ instead of f and using $h(r(\gamma))\, h(s(\gamma)) = 0$ for all $\gamma \in V$, we get $\varphi(f * b^2) = 0$ and $\varphi(f) = 0$, using a partition of unity. □

LEMMA 3.131. *Let φ be a KMS$_\beta$ state on $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$. Then, for any ideal $J \subset \mathcal{O}$ one has*

$$\varphi(e_J) = \mathfrak{n}(J)^{-\beta}\,.$$

PROOF. For each ideal J we let $\mu_J \in \mathcal{A}_{1,\mathbb{K}}$ be given as above by (3.469)

$$\mu_J((\Lambda, \phi), (\Lambda', \phi')) = \begin{cases} 1 & \Lambda = J^{-1}\Lambda' \text{ and } \phi' = \phi \\ 0 & \text{otherwise.} \end{cases}$$

One has $\sigma_t(\mu_J) = \mathfrak{n}(J)^{it}\, \mu_J$ for all $t \in \mathbb{R}$, while $\mu_J^*\, \mu_J = 1$ and $\mu_J\, \mu_J^* = e_J$. Thus the result just follows from the KMS condition. □

Given Proposition 3.115 above, we make the following definition.

DEFINITION 3.132. *A $\mathbb{K}$-lattice is* quasi-invertible *if the ρ in Proposition 3.115 is in $\hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$.*

Then we have the following result.

LEMMA 3.133. (1) *A $\mathbb{K}$-lattice (Λ, ϕ) that is divisible by only finitely many ideals is either quasi-invertible, or there is a finite place v such that the component ϕ_v of ϕ satisfies $\phi_v = 0$.*

(2) *A quasi-invertible* $\mathbb{K}$*-lattice is commensurable to a unique invertible* $\mathbb{K}$*-lattice.*

PROOF. Let (ρ, s) be associated to the $\mathbb{K}$-lattice (Λ, ϕ) as in Proposition 3.115. If $\rho \notin \mathbb{A}^*_{\mathbb{K},f}$, then either there exists a place v such that $\rho_v = 0$, or $\rho_v \neq 0$ for all v and there exists infinitely many places w such that $\rho_w^{-1} \notin \mathcal{O}_w$, where $\mathcal{O}_w$ is the local ring at w. This shows that the $\mathbb{K}$-lattice is divisible by infinitely many ideals. For the second statement, if we have $\rho \in \mathbb{A}^*_{\mathbb{K},f}$, we can write it as a product $\rho = s'_f \rho'$ where $\rho' = 1$ and $s'_f = \rho$. The $\mathbb{K}$-lattice obtained in this way is commensurable to the given one by Proposition 3.116 and is invertible. $\square$

We now complete the proof of the first two statements in Theorem 3.127.
Let φ be a KMS$_\beta$ state. Proposition 3.130 shows that there is a probability measure μ on $\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$ such that

$$\varphi(f) = \int_{\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)} f(L, L)\, d\mu(L), \quad \forall f \in \mathcal{A}_{1,\mathbb{K}}\,.$$

With $L = (\Lambda, \phi) = (\rho, s) \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$, Lemma 3.131 shows that the probability $\varphi(e_J)$ that an ideal J divides L is $\mathfrak{n}(J)^{-\beta}$. Since the series $\sum \mathfrak{n}(J)^{-\beta}$ converges for $\beta > 1$, it follows (cf. [**257**] Thm. 1.41) that, for almost all $L \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$, L is only divisible by a finite number of ideals. Notice that the KMS condition implies that the measure defined above gives measure zero to the set of $\mathbb{K}$-lattices (Λ, ϕ) such that the component $\phi_v = 0$ for some finite place v.
By the first part of Lemma 3.133, the measure μ gives measure one to quasi-invertible $\mathbb{K}$-lattices,

$$\mu\left((\hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}) \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)\right) = 1.$$

Notice that these $\mathbb{K}$-lattices form a Borel subset which is not closed. Then, by the second part of Lemma 3.133, the KMS$_\beta$ condition shows that the measure μ is entirely determined by its restriction to invertible $\mathbb{K}$-lattices, so that, for some probability measure ν,

$$\varphi = \int \varphi_{\beta,L}\, d\nu(L).$$

It follows that the Choquet simplex of extremal KMS$_\beta$ states is the space of probability measures on the compact space $\hat{\mathcal{O}}^* \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*) = \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ of invertible $\mathbb{K}$-lattices modulo scaling and its extreme points are the $\varphi_{\beta,L}$.

The action of symmetries $C_{\mathbb{K}}/D_{\mathbb{K}}$ on $\mathcal{E}_\beta$ is then free and transitive, as one can see from the fact that symmetries of $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ are induced from symmetries of $(\mathcal{A}_2, \sigma_t)$ and the action of $\mathrm{GL}_2^+(\mathbb{Q})$ on 2-dimensional $\mathbb{Q}$-lattices has only fixed points at those $\mathbb{Q}$-lattices (Λ, ϕ) with $\phi = 0$, as shown in Lemma 3.45.

Let $L = (\Lambda, \phi)$ be an invertible $\mathbb{K}$-lattice. For $s \in \hat{\mathcal{O}}^*/\mathcal{O}^* \subset \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$, the action is given explicitly by

$$\begin{aligned}(\varphi_{\beta,L} \circ \theta_s)(f) &= Z_{\mathbb{K}}(\beta)^{-1} \sum_{J \in \mathcal{J}_{\mathcal{O}}} f((J^{-1}\Lambda, s^{-1}\phi), (J^{-1}\Lambda, s^{-1}\phi))\, \mathfrak{n}(J)^{-\beta} \\ &= \varphi_{\beta,(\Lambda, s^{-1}\phi)}(f), \end{aligned} \tag{3.480}$$

for $f \in \mathcal{A}_{1,\mathbb{K}}$. More generally, for $s \in \hat{\mathcal{O}} \cap \mathbb{A}^*_{\mathbb{K},f}$, let $J_s = s\hat{\mathcal{O}} \cap \mathbb{K}$. One has

$$\begin{aligned}
(\varphi_{\beta,L} \circ \theta_s)(f) &= Z_{\mathbb{K}}(\beta)^{-1} \sum_{J \in \mathcal{J}_{\mathcal{O}}} \theta_s(f)((J^{-1}\Lambda, \phi), (J^{-1}\Lambda, \phi))\, \mathfrak{n}(J)^{-\beta} \\
&= Z_{\mathbb{K}}(\beta)^{-1} \sum_{J \subset J_s} f((J^{-1}\Lambda, s^{-1}\phi), (J^{-1}\Lambda, s^{-1}\phi))\, \mathfrak{n}(J)^{-\beta} \qquad (3.481) \\
&= Z_{\mathbb{K}}(\beta)^{-1} \sum_{J \in \mathcal{J}_{\mathcal{O}}} f(J^{-1}L_s, J^{-1}L_s)\, \mathfrak{n}(JJ_s)^{-\beta} \\
&= \mathfrak{n}(J_s)^{-\beta}\, \varphi_{\beta,L_s}(f),
\end{aligned}$$

for $f \in \mathcal{A}_{1,\mathbb{K}}$, and with L_s the invertible $\mathbb{K}$-lattice $(J_s^{-1}\Lambda, s^{-1}\phi)$. To prove the last equality one uses the basic property of Dedekind rings that any ideal $J \subset J_s$ can be written as a product of ideals $J = J'J_s$.

This completes the proof of the first two statements of the theorem. Now we prove the remaining statements of Theorem 3.127 on the extremal KMS states at zero-temperature.

The weak limits as $\beta \to \infty$ of states in $\mathcal{E}_\beta$ define states in $\mathcal{E}_\infty$ of the form

$$\varphi_{\infty,L}(f) = f(L, L). \qquad (3.482)$$

The action of the symmetry group $\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ on extremal KMS states at zero-temperature is given, as in §2.4, through the warming up/cooling down process described in (3.27).

In fact, for an invertible K-lattice L, evaluating $\varphi_{\infty,L}$ on $\theta_s(f)$ would not give a nontrivial action, while using (3.27) one obtains the action

$$\Theta_s(\varphi_{\infty,L})(f) = \lim_{\beta\to\infty} \left(W_\beta(\varphi_{\infty,L}) \circ \theta_s\right)(f), \qquad (3.483)$$

with W_β as in (3.27). This gives

$$\Theta_s(\varphi_{\infty,L}) = \varphi_{\infty,L_s}, \qquad (3.484)$$

with L_s as in (3.481).

Thus the action of the symmetry group $\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ is given by

$$L = (\Lambda, \phi) \mapsto L_s = (J_s^{-1}\Lambda, s^{-1}\phi), \quad \forall s \in \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*\,. \qquad (3.485)$$

When we evaluate states $\varphi_{\infty,L}$ on elements $f \in \mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$ of the arithmetic subalgebra we obtain

$$\varphi_{\infty,L}(f) = f(L, L) = g(L)\,, \qquad (3.486)$$

where the function g is the lattice function of weight 0 obtained as the restriction of f to the diagonal. By construction of the arithmetic subalgebra, one obtains in this way all the evaluations $f \mapsto f(z)$ of elements of the modular field F on the finitely many modules $z \in \mathbb{H}$ of the classes of $\mathbb{K}$-lattices.

Consider the subring B_τ of the modular field F consisting of those modular functions $f \in F$ that are defined at τ, where $\tau \in \mathbb{H}$ generates the imaginary quadratic field $\mathbb{K} = \mathbb{Q}(\tau)$. The theory of complex multiplication (cf. [**268**], §5) shows that the subfield $F_\tau \subset \mathbb{C}$ generated by the values $f(\tau)$, for $f \in B_\tau$, is the maximal abelian extension of $\mathbb{K}$ (where we have fixed an embedding $\mathbb{K} \subset \mathbb{C}$),

$$F_\tau = \mathbb{K}^{\mathrm{ab}}. \qquad (3.487)$$

Moreover, the action of $\alpha \in \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K})$ on the values $f(z)$ is given by

$$\alpha f(\tau) = f^{\,\tau q_\tau \theta^{-1}(\alpha^{-1})}(\tau). \tag{3.488}$$

This is Shimura reciprocity (cf. [**268**] Theorem 6.31 §6.8). In this formula, the notation $f \mapsto f^\gamma$ denotes the action of an element $\gamma \in \mathrm{Aut}(F)$ on the elements $f \in F$. The map θ is the class field theory isomorphism (3.127),

$$\theta : \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^* \to \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K}).$$

The map $q_\tau : \mathbb{A}^*_{\mathbb{K},f} \hookrightarrow \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ is the embedding determined by Proposition 3.113. The map τ is as in [**268**] Theorem 6.31, with the diagram with exact rows

$$\begin{tikzcd} 1 \arrow[r] & \mathbb{K}^* \arrow[r, "\iota"] & \mathrm{GL}_1(\mathbb{A}_{\mathbb{K},f}) \arrow[r, "\theta"] \arrow[d, "q_\tau"] & \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K}) \arrow[r] & 1 \\ 1 \arrow[r] & \mathbb{Q}^* \arrow[r] & \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \arrow[r, "\tau"] & \mathrm{Aut}(F) \arrow[r] & 1. \end{tikzcd} \tag{3.489}$$

This result makes essential use of an important result of Shimura (cf. [**268**] §6.6) recalled above in Theorem 3.75 that characterizes the automorphism group of the modular field by the exact sequence

$$1 \to \mathbb{Q}^* \to \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \to \mathrm{Aut}(F) \to 1.$$

In fact, using the notations of Theorem 3.75, the above equation (3.488) becomes

$$\alpha f(\tau) = (\mathrm{Gal}(q'_\tau(\theta^{-1}(\alpha^{-1})))f)(\tau), \tag{3.490}$$

where we used (3.305) and Proposition 3.113 to replace q_τ by q'_τ.
Thus, when we act by an element $\alpha \in \mathrm{Gal}(\mathbb{K}^{\mathrm{ab}}/\mathbb{K})$ on the values on $\mathcal{A}_{1,\mathbb{K},\mathbb{Q}} \otimes \mathbb{K}$ of an extremal KMS_∞ state we have

$$\alpha\, \varphi_{\infty,L}(f) = \mathrm{Gal}_\tau(q'_\tau(\theta^{-1}(\alpha^{-1})))\varphi_{\infty,L}(f)$$

and using Theorem 3.100

$$\alpha\, \varphi_{\infty,L}(f) = \varphi_{\infty,L} \circ \vartheta_g(f)\,, \quad g = q'_\tau(\theta^{-1}(\alpha^{-1}))\,.$$

Thus, using the relation (3.473) between the symmetry groups, we get

$$\alpha\, \varphi_{\infty,L}(f) = \varphi_{\infty,L_s}(f), \tag{3.491}$$

where $s = \theta^{-1}(\alpha) \in \mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$.
This completes the result on the Galois properties of zero temperature KMS states. $\square$

8.9. High temperature KMS states.

The main result on the behavior of the system at high temperature (i.e. for $\beta \leq 1$) is given by the following ([**90**]).

THEOREM 3.134. *The system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ undergoes a phase transition with spontaneous symmetry breaking at inverse temperature $\beta = 1$. In the range $0 < \beta \leq 1$ there is a unique extremal KMS state.*

Proof. The proof follows along the lines of [**236**]. We first discuss uniqueness. By Proposition 3.130, one obtains a measure μ on the space

$$X = \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} (\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*)$$

of $\mathbb{K}$-lattices up to scale. As in Lemma 3.96, this measure fulfills the quasi-invariance condition

$$\int_X \mu_J \, f \, \mu_J^* \, d\mu = \mathfrak{n}(J)^{-\beta} \int_X f \, d\mu, \tag{3.492}$$

for all ideals J, where μ_J is as in (3.469). To prove uniqueness of such a measure, for $\beta \in (0,1]$, one proceeds in the same way as in [**236**], reducing the whole argument to an explicit formula for the orthogonal projection P from $L^2(X, d\mu)$ to the subspace of functions invariant under the semigroup action

$$L = (\Lambda, \phi) \mapsto J^{-1}L, \tag{3.493}$$

which preserves commensurability. As in [**236**], one can obtain such a formula as a weak limit of the orthogonal projections P_A associated to finite sets A of non-Archimedean places.

Let A be a finite set of non-Archimedean places. Let $\mathcal{J}_A$ be the subsemigroup of the semigroup $\mathcal{J}_{\mathcal{O}}$ of ideals generated by the prime ideals in A. Any element $J \in \mathcal{J}_A$ can be uniquely written as a product

$$J = \prod_{v \in A} J_v^{n_v}, \tag{3.494}$$

where J_v is the prime ideal associated to the place $v \in A$.

Lemma 3.135. *Let $L = (\Lambda, \phi)$ be a $\mathbb{K}$-lattice such that $\phi_v \neq 0$ for all $v \in A$. Let $J \in \mathcal{J}_A$, $J = \prod_{v \in A} J_v^{n_v}$ be the smallest ideal dividing L. Let $(\rho, s) \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} \mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*$ be the pair associated to L. Then, for each $v \in A$, the valuation of ρ_v is equal to n_v.*

Proof. Let (ρ, s) be as above, and m_v be the valuation of ρ_v. Then it is enough to show that an ideal J divides L if and only if J is of the form (3.494), with $n_v \leq m_v$. The map ϕ is the composite of multiplication by ρ and an isomorphism, as in the diagram (3.452), hence the divisibility is determined by the valuations of ρ_v. □

Definition 3.136. *With A as above we shall say that a $\mathbb{K}$-lattice $L = (\Lambda, \phi)$ is A-invertible if and only if the valuation of ρ_v is equal to zero for all $v \in A$.*

We now define basic test functions associated to a Hecke Grössencharakter. Given such a character χ, the restriction of χ to $\hat{\mathcal{O}}^*$ only depends on the projection on $\hat{\mathcal{O}}^*_{B_\chi} = \prod_{v \in B_\chi} \hat{\mathcal{O}}^*_v$, for B_χ a finite set of non-Archimedean places. Let B be a finite set of non-Archimedean places $B \supset B_\chi$. We consider the function $f = f_{B,\chi}$ on $\hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} \mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*$, which is obtained as follows. For $(\rho, s) \in \hat{\mathcal{O}} \times_{\hat{\mathcal{O}}^*} \mathbb{A}^*_{\mathbb{K}}/\mathbb{K}^*$, we let $f = 0$ unless $\rho_v \in \hat{\mathcal{O}}^*_v$ for all $v \in B$, while $f(\rho, s) = \chi(\rho' s)$, for any $\rho' \in \hat{\mathcal{O}}^*$ such that $\rho'_v = \rho_v$ for all $v \in B$. This is well defined, because the ambiguity in the choice of ρ' does not affect the value of χ, since $B_\chi \subset B$. The function f obtained this way is continuous.

Let $H(B)$ be the subspace of $L^2(X, d\mu)$ of functions that only depend on s and on the projection of ρ on $\hat{\mathcal{O}}_B$. Let us consider the map that assigns to a $\mathbb{K}$-lattice L the smallest ideal $J \in \mathcal{J}_B$ dividing L, extended by zero if some $\phi_v = 0$. By Lemma 3.135, the value of this map only depends on the projection of ρ on $\hat{\mathcal{O}}_B$. Thus, we can consider corresponding projections $E_{B,J}$ in $H(B)$, for J as above. By construction the projections $E_{B,J}$ give a partition of unity on the Hilbert space $H(B)$. Note that $E_{B,0} = 0$, since the measure μ gives measure zero to the set of $\mathbb{K}$-lattices with $\phi_v = 0$ for some v.

Let $V_J f(L) = f(J^{-1}L)$ implementing the semigroup action (3.493). For $J \in \mathcal{J}_B$, the operator $\mathfrak{n}(J)^{-\beta/2} V_J^*$ maps isometrically the range of $E_{B,\mathcal{O}}$ to the range of $E_{B,J}$.

LEMMA 3.137. *The functions $V_J^* f_{B,\chi}$ span a dense subspace of $H(B)$.*

PROOF. It is sufficient to prove that the $f_{B,\chi}$ form a dense subspace of the range of $E_{B,\mathcal{O}}$. The image of $\hat{\mathcal{O}}^*$ in $\hat{\mathcal{O}}_B^* \times \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$ by the map $u \mapsto (u, u^{-1})$ is a closed normal subgroup. We let $\hat{\mathcal{O}}_B^* \times_{\hat{\mathcal{O}}^*} \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$ be the quotient. This is a locally compact group. The quotient G_B by the connected component of identity $D_{\mathbb{K}}$ in $\mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$ is a compact group. Then $C(G_B)$ is identified with a dense subspace of the range of $E_{B,\mathcal{O}}$. The characters of G_B are the Grossencharakters χ that vanish on the connected component of the identity and such that $B_\chi \subset B$. Thus, by Fourier transform, we obtain the density result. □

Let A be a finite set of non-Archimedean places, and $\mathcal{J}_A$ as above. Let H_A be the subspace of functions constant on $\mathcal{J}_A$-orbits, and let P_A be the corresponding orthogonal projection. The P_A converge weakly to P.

PROPOSITION 3.138. *Let $A \supset B$ be finite sets of non-Archimedean places. Let L be an A-invertible $\mathbb{K}$-lattice, and $f \in H(B)$. The restriction of $P_A f$ to the $\mathcal{J}_A$-orbit of L is constant and given by the formula*

$$P_A f|_{\mathcal{J}_A L} = \zeta_{\mathbb{K},A}(\beta)^{-1} \sum_{J \in \mathcal{J}_A} \mathfrak{n}(J)^{-\beta} f(J^{-1}L), \tag{3.495}$$

where $\zeta_{\mathbb{K},A}(\beta) = \sum_{J \in \mathcal{J}_A} \mathfrak{n}(J)^{-\beta}$.

PROOF. By construction, the right hand side of the formula (3.495) defines an element f_A in $H_A \cap H(A)$. One checks, using the quasi-invariance condition (3.492) on the measure μ, that $\langle f_A, g\rangle = \langle f, g\rangle$ for all $g \in H_A$, as in [**236**]. □

Let L be an invertible $\mathbb{K}$-lattice and χ a Grössencharakter vanishing on the connected component of the identity $D_{\mathbb{K}}$. We define $\chi(L)$ as $\chi(\rho s)$, for any representative (ρ, s) of L. This continues to make sense when L is an A-invertible $\mathbb{K}$-lattice and $A \supset B_\chi$ taking $\chi(\rho' s)$ where $\rho' \in \hat{\mathcal{O}}^*$ and $\rho'_v = \rho_v$ for all $v \in A$.

Finally, we recall that to a Grössencharakter χ vanishing on the connected component of the identity $D_{\mathbb{K}}$ one associates a Dirichlet character $\tilde{\chi}$ defined for ideals J in $\mathcal{J}_{B_\chi^c}$, where B_χ^c is the complement of B_χ. More precisely, given $J \in \mathcal{J}_{B_\chi^c}$, let s_J be an idèle such that $J = s_J \hat{\mathcal{O}} \cap \mathbb{K}$ and $(s_J)_v = 1$ for all places $v \in B_\chi$. One then define $\tilde{\chi}(J)$ to be the value $\chi(s_J)$. This is independent of the choice of such s_J.

PROPOSITION 3.139. *Let $A \supset B \supset B_\chi$ and L an A-invertible $\mathbb{K}$-lattice. The projection P_A of* (3.495) *applied to the function $f_{B,\chi}$ gives*

$$P_A f_{B,\chi}|_{\mathcal{J}_A L} = \frac{\chi(L)}{\zeta_{\mathbb{K},A}(\beta)} \sum_{J \in \mathcal{J}_{A \setminus B}} \mathfrak{n}(J)^{-\beta}\, \tilde{\chi}(J)^{-1}. \tag{3.496}$$

PROOF. Among ideals in $\mathcal{J}_A$, those that have nontrivial components on B do not contribute to the sum (3.495) computing $P_A f_{B,\chi}|_{\mathcal{J}_A L}$. It remains to show that $f_{B,\chi}(J^{-1}L) = \chi(L)\tilde{\chi}(J)^{-1}$, for $J \in \mathcal{J}_{A\setminus B}$. Let $J = s_J \hat{\mathcal{O}} \cap \mathbb{K}$ and $(s_J)_v = 1$ for all places $v \in B$. Let L be given by (ρ, s). We have $J^{-1}L$ given by $(\rho\, s_J, s\, s_J^{-1})$ using Proposition 3.122. Thus, for any choice of $\rho' \in \hat{\mathcal{O}}^*$ with $\rho'_v = (\rho s_J)_v$ for all $v \in B$ one has $f_{B,\chi}(J^{-1}L) = \chi(\rho' s s_J^{-1}) = \chi(\rho' s)\tilde{\chi}(J)^{-1}$. Note that $(s_J)_v = 1$ for all places $v \in B$. Thus the choice of ρ' is governed by $\rho'_v = \rho_v$ for all $v \in B$. Since L is A-invertible and $A \supset B \supset B_\chi$ we get $\chi(\rho' s) = \chi(L)$ for a suitable choice of ρ'. □

It then follows as in [**236**] that the $P_A f_{B,\chi}$ tend weakly to zero for χ nontrivial. The same argument gives an explicit formula for the measure, obtained as a limit of the $P_A f_{B,1}$, for the trivial character. In particular, the restriction of the measure to G_B is proportional to the Haar measure. Positivity is ensured by the fact that we are taking a projective limit of positive measures. This completes the proof of existence and uniqueness of the KMS_β state for $\beta \in (0, 1]$. □

Notice that the structure of KMS states for the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ differs substantially from that of the system $(\mathcal{A}_2, \sigma_t)$, even though $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ is constructed as a subsystem of $(\mathcal{A}_2, \sigma_t)$.

So, the system $(\mathcal{A}_2, \sigma_t)$ has two phase transitions at $\beta = 1$ and $\beta = 2$ with no KMS states at all in the high temperature regime $0 < \beta < 1$, a unique KMS state in the range $1 < \beta \leq 2$ and a large set of extremal KMS states given by the points of the Shimura variety of GL_2. On the other hand the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ has, like the Bost–Connes system, a unique phase tansition at $\beta = 1$ with a unique KMS_β state in the range $0 < \beta \leq 1$ and extremal KMS states given by points of $\mathbb{A}^*_{\mathbb{K},f}/\mathbb{K}^*$ for $1 < \beta \leq \infty$.

This different behavior is not surprising, as the following general fact illustrates.

Given an étale groupoid $\mathcal{G}$ and a full subgroupoid $\mathcal{G}' \subset \mathcal{G}$, let ρ be a homomorphism $\rho : \mathcal{G} \to \mathbb{R}^*_+$. The inclusion $\mathcal{G}' \subset \mathcal{G}$ gives a correspondence between the C^*-algebras associated to $\mathcal{G}'$ and $\mathcal{G}$, compatible with the time evolution associated to ρ and its restriction to $\mathcal{G}'$.

The following simple example, however, shows that, in general, the KMS states for the $\mathcal{G}'$ system do not map to KMS states for the $\mathcal{G}$ system.

We let $\mathcal{G}$ be the groupoid with units $\mathcal{G}^{(0)}$ given by an infinite countable set, and morphisms given by all pairs of units. Consider a finite subset of $\mathcal{G}^{(0)}$ and let $\mathcal{G}'$ be the reduced groupoid. Finally, let ρ be trivial. Clearly, the $\mathcal{G}'$ system admits a KMS state for all temperatures given by the trace, while, since there is no tracial state on the compact operators, the $\mathcal{G}$ system has no KMS states.

8.10. Comparison with other systems.

We have seen in (3.455) that the groupoid $\mathcal{G}_{1,\mathbb{K}}$ corresponds to the quotient $\mathbb{A}^{\cdot}_{\mathbb{K}}/\mathbb{K}^*$ viewed as a noncommutative space. Thus, its C^*-algebra is given up to Morita equivalence by the crossed product $C_0(\mathbb{A}^{\cdot}_{\mathbb{K}}) \rtimes \mathbb{K}^*$ as in (3.455). In the case of the groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$, we are similarly looking at the quotient $\mathbb{A}_{\mathbb{K},f}/\mathbb{K}^*$, with the noncommutative algebra of coordinates, which, up to Morita equivalence, can be identified with the crossed product

$$C_0(\mathbb{A}_{\mathbb{K},f}/\mathcal{O}^*) \rtimes \mathbb{K}^*/\mathcal{O}^*. \tag{3.497}$$

where $\mathcal{O}^*$ is the finite group of units.
We are going to show more precisely the Morita equivalence between this algebra and the convolution algebra $\mathcal{A}_{1,\mathbb{K}} = C^*(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)$ discussed in the previous paragraphs. Let us compare the groupoid $\mathbb{A}_{\mathbb{K},f}/\mathcal{O}^* \rtimes \mathbb{K}^*/\mathcal{O}^*$ with the groupoid $\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*$. The difference can be seen by looking at the case of $\mathbb{K}$-lattices with $\phi = 0$. In the first case, this corresponds to the point $0 \in \mathbb{A}_{\mathbb{K},f}/\mathcal{O}^*$, which has stabilizer $\mathbb{K}^*/\mathcal{O}^*$. Thus, we obtain the group C^*-algebra of $\mathbb{K}^*/\mathcal{O}^*$. In the other case, the corresponding groupoid is obtained as a quotient by $\mathbb{C}^*$ of the groupoid $\mathcal{G}_{1,\mathbb{K},0} \subset \mathcal{G}_{1,\mathbb{K}}$ of pairs of commensurable $\mathcal{O}$-modules (finitely generated of rank one) in $\mathbb{C}$. In this case the units of the groupoid $\mathcal{G}_{1,\mathbb{K},0}/\mathbb{C}^*$ can be identified with the elements of $\mathrm{Cl}(\mathcal{O})$. These elements are pairwise commensurable and the reduced groupoid by any of these units is the group $\mathbb{K}^*/\mathcal{O}^*$ so that one obtains the required Morita equivalence on that part. The result below gives the Morita equivalence in general.

PROPOSITION 3.140. *Let $\mathcal{A}_{\mathbb{K},princ} := C_0(\mathbb{A}_{\mathbb{K},f}/\mathcal{O}^*) \rtimes \mathbb{K}^*/\mathcal{O}^*$. Let $\mathcal{H}$ be the space of pairs up to scaling $((\Lambda,\phi),(\Lambda',\phi'))$ of commensurable $\mathbb{K}$-lattices, with (Λ,ϕ) principal. The space $\mathcal{H}'$ is defined analogously, with (Λ',ϕ') principal. Then $\mathcal{H}$ (resp. $\mathcal{H}'$) has the structure of $\mathcal{A}_{\mathbb{K},princ}$–$\mathcal{A}_{1,\mathbb{K}}$ (resp. $\mathcal{A}_{1,\mathbb{K}}$–$\mathcal{A}_{\mathbb{K},princ}$) bimodule. These bimodules give a Morita equivalence between the algebras $\mathcal{A}_{\mathbb{K},princ}$ and $\mathcal{A}_{1,\mathbb{K}} = C^*(\mathcal{G}_{1,\mathbb{K}}/\mathbb{C}^*)$.*

PROOF. The correspondence given by these bimodules has the effect of reducing to the principal case. In that case the groupoid of the equivalence relation (not up to scaling) is given by the crossed product $\mathbb{A}^{\cdot}_{\mathbb{K}} \rtimes \mathbb{K}^*$. When taking the quotient by $\mathbb{C}^*$ we then obtain the groupoid $\mathbb{A}_{\mathbb{K},f}/\mathcal{O}^* \rtimes \mathbb{K}^*/\mathcal{O}^*$. □

In particular, this result gives an explicit Morita equivalence between the algebra $\mathcal{A}_{1,\mathbb{K}}$ considered here and the algebra considered in [**194**].

Proposition 3.140 also gives the functoriality between the systems $(\mathcal{A}_{1,\mathbb{K}},\sigma)$ for imaginary quadratic fields and the Bost–Connes system $(\mathcal{A}_1,\sigma_t)$ for $\mathbb{Q}$, which corresponds to the field extension.
By Proposition 3.140, we can replace the C^*-algebra $\mathcal{A}_{1,\mathbb{K}}$ by the Morita equivalent $\mathcal{A}_{\mathbb{K},princ} = C_0(\mathbb{A}_{\mathbb{K},f}/\mathcal{O}^*) \rtimes \mathbb{K}^*/\mathcal{O}^*$. We need to relate it to the BC system, which in turn is described up to Morita equivalence, using the crossed product algebra $C_0(\mathbb{A}_{\mathbb{Q},f}) \rtimes \mathrm{GL}_1^+(\mathbb{Q})$.
Let $\mathfrak{n} : \mathbb{A}_{\mathbb{K},f} \to \mathbb{A}_{\mathbb{Q},f}$ be the norm map. Its range $\mathfrak{n}(\mathbb{K}^*)$ is a subgroup of $\mathrm{GL}_1^+(\mathbb{Q})$ and, with $\mathbb{K} = \mathbb{Q}(\sqrt{-d})$, d a square-free positive integer, is determined by the conditions

$$\rho \in \mathfrak{n}(\mathbb{K}^*) \Leftrightarrow (\rho,-d)_p = 1\,, \quad \forall\,\text{prime}\;\; p \tag{3.498}$$

where $(a,b)_p$ denotes the Hilbert symbol ([**262**]). One checks that $\mathfrak{n}(\mathbb{K}^*)$ is a free abelian group. Let then $\iota : \mathfrak{n}(\mathbb{K}^*) \to \mathbb{K}^*$ be a group homomorphism that gives a section of $\mathfrak{n} : \mathbb{K}^* \to \mathfrak{n}(\mathbb{K}^*)$.

PROPOSITION 3.141. *The pair $(\mathfrak{n},\iota)$ induces a morphism*

$$\varrho : C_0(\mathbb{A}_{\mathbb{Q},f}) \rtimes \mathfrak{n}(\mathbb{K}^*) \to M(C_0(\mathbb{A}_{\mathbb{K},f}/\mathcal{O}^*) \rtimes \mathbb{K}^*/\mathcal{O}^*) = M(\mathcal{A}_{\mathbb{K},princ})$$

to the multiplier algebra of $\mathcal{A}_{\mathbb{K},princ}$.

PROOF. The argument is analogous to the one used in proving the functoriality of §7.10. One checks that the condition (3.434) is fulfilled by construction. Note that, in the setup as above, $\mathfrak{n}(\mathbb{K}^*)$ is a subgroup of $\mathrm{GL}_1^+(\mathbb{Q})$ and, moreover, $\mathfrak{n}(\mathcal{O}^*) = 1$.

The noncommutative space $\mathbb{A}_{\mathbb{Q},f}/\mathfrak{n}(\mathbb{K}^*)$, obtained as quotient of $\mathbb{A}_{\mathbb{Q},f}$ by $\mathfrak{n}(\mathbb{K}^*)$, is an etale infinite covering of the space of one-dimensional $\mathbb{Q}$-lattices up to scaling. $\square$

9. Quantum statistical mechanics of Shimura varieties

As we have seen, the system $(\mathcal{A}_2, \sigma_t)$ is closely related to the geometric and arithmetic structure of modular curves. The modular curves are obtained as quotients $\Gamma' \backslash \mathbb{H}$ of the uper half plane by a finite index congruence subgroup $\Gamma' \subset \mathrm{SL}_2(\mathbb{Z})$, such as the congruence subgroups $\Gamma(N)$. They form a tower of branched coverings as one varies the subgroup. The projective limit of this tower sits as a connected component in the *adelic* quotient

$$\mathrm{GL}_2(\mathbb{Q})\backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}})/\mathbb{C}^*. \tag{3.499}$$

An important property of the space (3.499) is the fact that it is a moduli space. In fact, the space (3.499) along with its zero-dimensional space of components

$$\mathrm{GL}_1(\mathbb{Q})\backslash \mathrm{GL}_1(\mathbb{A}_{\mathbb{Q}})/\mathbb{R}_+^* = \pi_0(\mathrm{GL}_2(\mathbb{Q})\backslash \mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}})/\mathbb{C}^*) \tag{3.500}$$

underlying the Bost–Connes system belong to an important class of algebro-geometric moduli spaces, the Shimura varieties. These are moduli spaces of Hodge structures (cf. e.g. [**228**]).

Roughly speaking (see [**112**] for a detailed treatment of the subject) a Shimura variety $Sh(G, X)$ is constructed from data (G, X) of a reductive algebraic group G and a Hermitian symmetric domain X. For example, (3.499) is the Shimura variety $Sh(\mathrm{GL}_2, \mathbb{H}^{\pm})$, where $\mathbb{H}^{\pm} = \mathrm{GL}_2(\mathbb{R})/\mathbb{C}^*$ is the union of the upper and the lower half plane in $\mathbb{P}^1(\mathbb{C})$, while (3.500) is the Shimura variety $Sh(\mathrm{GL}_1, \{\pm\})$, as we already discussed in previous sections of this chapter.

In both systems $(\mathcal{A}_1, \sigma_t)$ and $(\mathcal{A}_2, \sigma_t)$ the noncommutative space is obtained by replacing the adelic points $G(\mathbb{A}_{\mathbb{Q},f})$ of the reductive group G in the Shimura variety with the adelic points $M(\mathbb{A}_{\mathbb{Q},f})$ of an associated semigroup M, such that $G \subset M$ is the set of invertible elements. In the case of $(\mathcal{A}_1, \sigma_t)$ this is obtained by replacing $\mathbb{A}^*_{\mathbb{Q},f}$ by $\mathbb{A}_{\mathbb{Q},f}$, while in the case of $(\mathcal{A}_2, \sigma_t)$ one replaces $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f})$ with 2×2 matrices $M_2(\mathbb{A}_{\mathbb{Q},f})$.

This observation was adopted in the recent work of Ha and Paugam [**159**] to systematically produce a large collection of systems $(\mathcal{A}, \sigma_t)$ associated to Shimura varieties $Sh(G, X)$. These generalize and unify all the cases treated in this chapter (the Bost–Connes system, the GL_2-system, and the systems associated to imaginary quadratic fields) within the same general setting.

For a Shimura variety $Sh(G, X)$, given a faithful representation $\rho : G \to \mathrm{GL}(V)$ as above, there is an "enveloping semigroup" M, that is, a normal irreducible semigroup $M \subset \mathrm{End}(V)$ such that $M^* = G$. The data (G, X, V, M) then determine a noncommutative space

$$Sh_K^{\mathrm{nc}}(G, X, V, M) = G(\mathbb{Q})\backslash M(\mathbb{A}_f) \times X\,.$$

(This covers the case where $Sh(G, X) = G(\mathbb{Q})\backslash G(\mathbb{A}_{\mathbb{Q},f}) \times X$ but more general cases are also covered by the results of [**159**].) The resulting noncommutative space can still be given a modular interpretation in terms of degenerations of the Hodge data. The noncommutative space $Sh_K^{\mathrm{nc}}(G, X, V, M)$ can be described by a convolution algebra. This involves some careful use of the theory of stacks to treat the presence of singularities (a phenomenon we already saw in the case of the GL_2-system, which

pervades systematically the higher dimensional cases). The resulting convolution algebra $\mathcal{A}_{G,X,V,M}$ has a natural time evolution σ_t, so that one can associate to a Shimura variety a quantum statistical mechanical system $(\mathcal{A}_{G,X,V,M}, \sigma_t)$. These systems have partition functions that define a new class of zeta functions associated to Shimura varieties. The points of the classical Shimura variety $Sh(G, X)$ appear as extremal low temperature KMS states of $(\mathcal{A}_{G,X,V,M}, \sigma_t)$.
In particular, as part of this general construction, the work of [**159**] also provides a natural generalization of the Bost–Connes system $(\mathcal{A}_1, \sigma_t)$ to arbitrary number fields. This agrees with the system $(\mathcal{A}_{1,\mathbb{K}}, \sigma_t)$ of [**90**] in the case of imaginary quadratic fields. The general system for a number field $\mathbb{K}$ has partition function the Dedekind zeta function and symmetries given by the group $C_{\mathbb{K}}/D_{\mathbb{K}}$.

CHAPTER 4

Endomotives, thermodynamics, and the Weil explicit formula

The spectral realization of the zeros of L-functions described in Chapter 2 made little explicit use of the formalism of noncommutative geometry, except for the use of the quantized calculus in proving the semi-local trace formula.
One of our goals in this chapter is to clarify the conceptual meaning of the spectral realization and in particular of the map $\mathfrak{E}$ of Chapter 2 §6 in terms of noncommutative geometry and cyclic cohomology. This chapter is based on our joint work with Consani [**74**], [**75**].

There are three essential ingredients in the conceptual understanding of the spectral realization:

(1) Geometry
(2) Thermodynamics
(3) Cohomology and motives

In the first step we extend the notion of zero-dimensional motive, i.e. of Artin motive, to the noncommutative set-up. The main examples of a noncommutative Artin motive we are interested in arise from endomorphisms of algebraic varieties. This is in fact the origin of the terminology "endomotive" we use for these zero-dimensional noncommutative spaces. The reason for thinking of this class of noncommutative spaces as motives lies in the fact that we define morphisms as correspondences, generalizing the correspondences by algebraic cycles used in the theory of motives.
We work over an algebraic number field $\mathbb{K}$ and the absolute Galois group, i.e. the Galois group G of $\bar{\mathbb{K}}/\mathbb{K}$, will act on the various objects we construct (cf. §2.3). Given an abelian semigroup S of endomorphisms of an algebraic variety Y and a point $y_0 \in Y$ fixed under S, the projective limit X of the preimages of y_0 under elements of S admits a natural action of S and this gives rise to an endomotive (X, S). Such datum is best encoded as an algebraic groupoid $\mathcal{G} = X \rtimes S$ and the associated convolution algebra $\mathcal{A}_{\mathbb{K}}$.
We first recall in §1 the basic notions of morphisms for noncommutative spaces provided by Kasparov's KK-theory in the context of C^*-algebras (§1.1) and by the abelian category of cyclic modules in the general algebraic context (§§1.2, 1.3). We describe the relation of the cyclic category to cyclic (co)homology in §1.4.
We then present in §2 the construction of a category of *endomotives* that extends the notion of morphism given by geometric correspondences from the case of Artin motives to the noncommutative endomotives. We construct this category first at the purely algebraic level, i.e. dealing with zero-dimensional pro-varieties over $\mathbb{K}$, in §2.1. We then give a construction at the analytic level in §2.2, working with totally disconnected locally compact spaces. We show in §2.4 that under a "uniform" condition on the projective system that determines an endomotive, this comes endowed

with a probability measure, which induces a state on the corresponding C^*-algebra.

We then show in Theorem 4.34 how the functor $\mathcal{P}$ that replaces a variety over $\mathbb{K}$ by its set of $\bar{\mathbb{K}}$ points gives a natural relation between the categories of algebraic and analytic endomotives, cf. §2.5. Moreover, the Galois group G acts by natural transformations of $\mathcal{P}$.
In §2.6 we describe the construction of the endomotive associated to an abelian semigroup S of endomorphisms of an algebraic variety Y. We then show in §2.7 that the simplest example, i.e. the semi-group of endomorphisms of the multiplicative group $\mathbb{G}_m$ for $\mathbb{K} = \mathbb{Q}$, gives an endomotive that coincides at the analytic level with the Bost-Connes system analyzed in Chapter 3.
We discuss in §3 the problem of extending correspondences from endomotives to noncommutative spaces in higher dimension, in a way that is compatible with the definition of correspondences in the algebro-geometric context in terms of K-theory. In §3.1 we recall the setting of geometric correspondences in KK-theory. In §3.2 we compare this with the cycle map defined using algebraic K-theory and we comment in §3.2.1 on the relation between algebraic and topological K-theory and cyclic and Hochschild (co)homology. We also comment on the relation between motives and noncommutative spaces in the context of noncommutative tori and abelian varieties in §3.2.2.
As a conclusion of this first step of the construction, we obtain from the data of a uniform endomotive a noncommutative geometric datum given by a pair $(\mathcal{A}, \varphi)$ of an involutive algebra and a state, together with an action by automorphisms of the Galois group G.
The second step then involves thermodynamics, which we develop in §4. As will become apparent in this chapter, a basic new feature of noncommutative spaces plays a dominant role in our interpretation: these spaces have thermodynamical properties and in particular they can be analyzed at different temperatures by classifying the KMS_β states for the modular automorphism group σ_t^φ, as we already did in in Chapter 3 for three basic examples discussed there.
What we saw in Chapter 3 is that, by lowering the temperature, a given noncommutative space tends to become more and more commutative or *classical*, so that in good situations the extremal KMS_β states play the role of *classical points* of the space. The key feature of extremal KMS_β states that can play the role of points is that the corresponding factor (obtained from the state by the GNS construction) is Morita equivalent to a point, i.e. it is of type I.
After recalling in §4.1 Tomita's theory of the modular automorphism group σ_t^φ, we analyze the type I extremal KMS_β states in §4.2 and we show in particular (cf. Lemma 4.56) that such states persist at lower temperatures. Thus, cooling down the system by lowering the temperature has the effect of adding more and more classical points of the noncommutative space.
The space Ω of classical points comes naturally equipped with a principal $\mathbb{R}_+^*$-bundle $\tilde{\Omega}$. In Theorem 4.85 we obtain a conceptual understanding of the map $\mathfrak{E}$ of Chapter 2 §6 as the natural morphism π of restriction from the dual system defined in §4.3 to the $\mathbb{R}_+^*$-bundle $\tilde{\Omega}$ over the space of classical points, cf. §§4.5, 4.6, 4.7. We first deal with general systems in §4.8. We then specialize to the BC system in §4.9.
We also explain in §4.4 why passing to the dual system is the analogue in characteristic zero of the unramified extension $\mathbb{K} \otimes_{\mathbb{F}_q} \bar{\mathbb{F}}_q$ for a global field of positive

characteristic. This gives a first indication for the interpretation of the scaling action on the dual system as a characteristic zero analog of the action of Frobenius that we go on to develop in the subsequent sections.
The third step, in fact, provides a replacement in characteristic zero for the action of Frobenius étale cohomology, given in terms of the scaling action on the cyclic homology of the cokernel of the above restriction map π. This is described in §§4.8, 4.9, 4.10. The latter subsection relates the scaling action on the cyclic homology of the cokernel of the restriction map of the Bost–Connes endomotive to the spectral realization of zeros of the Riemann zeta function.
It is for this analysis of the scaling action and the correct definition of the cokernel of the restriction map that one needs to work in an abelian category of motivic nature, so that one can make sense of the cokernel of a morphism of algebras. This abelian category is the category of Λ-modules, the central tool of cyclic cohomology, which was introduced in §1.2. Since traces Tr define cyclic morphisms $\mathrm{Tr}^\natural$ one can compose the restriction morphism π with the trace on trace class operators and obtain, using the results of §4.5, a cyclic morphism $(\mathrm{Tr}\circ\pi)^\natural$ whose range is contained in the cyclic module of the *commutative* algebra $C(\tilde{\Omega})$. The cyclic homology of the cokernel produces a representation of the product $G\times\mathbb{R}_+^*$ of the Galois group by the multiplicative group $\mathbb{R}_+^*$. The main result is then to apply this general framework to a specific object in the category of endomotives, which corresponds to the Bost–Connes system, and obtain a cohomological version of the Weil explicit formula as a trace formula.
These results are then extended to an arbitrary global field $\mathbb{K}$ in §5. In §5.1 we describe in this generality the adèle class space of a global field and its algebra of coordinates as a noncommutative space. In §5.2 we describe the corresponding cyclic module in the abelian category of Λ-modules and in §5.3 we give the general form of the restriction map corresponding to the inclusion of the idèle class space as "classical points" of the noncommutative space of adèle classes, as a cyclic morphism.
We describe this more explicitly in the case of $\mathbb{K}=\mathbb{Q}$ in §5.4. In §5.5 we return to the general case of global fields and we show how the cokernel of the restriction map can be thought of as a motivic H^1 with an induced action of the idèle class group $C_{\mathbb{K}}$. We show that this gives the spectral realization of the zeros of L-functions with Grössencharakter. We show then in §5.6 that the action on this H^1 of functions in the "strong Schwartz space" of $C_{\mathbb{K}}$ that are in the range of the restriction map vanishes identically. Finally, in §5.7 we give the formulation of the Weil explicit formula as a Lefschetz trace formula for the action on the H^1 of §5.5 of elements in the "strong Schwartz space" of $C_{\mathbb{K}}$. We then show in §5.8 that, in terms of this Lefschetz trace formula, the Riemann Hypothesis becomes a statement equivalent to positivity of the induced trace pairing.
We then begin with §6 and §7 a comparative analysis of the Weil proof of the Riemann Hypothesis for function fields and the noncommutative geometry of the adèle class space, as in [**75**]. The aim is to develop a sufficiently rich dictionary that will eventually provide good noncommutative geometry analogs for all the main algebro-geometric notions involved in the Weil proof, such as algebraic curves, divisors, linear equivalence, Riemann–Roch theorem, and Weil positivity.
We begin in §6 by recalling the essential steps in the Weil proof of the Riemann Hypothesis for function fields.

We continue the chapter by drawing some compelling analogies between the Weil proof of the Riemann Hypothesis for function fields and the setting of noncommutative geometry.

In §7 we begin to build the corresponding noncommutative geometry notion, that will be summarized in the dictionary in §7.7. We begin by recalling in §7.1 the computation of the distributional trace of a flow on a smooth manifold. We then describe in §7.2 the periodic orbits of the action of $C_{\mathbb{K}}$ on the adèle class space.

This gives a natural definition of the scaling correspondence that parallels the role of the Frobenius correspondence in the Weil proof (cf. §7.3). We then show in §7.4 that the step in the Weil proof of adjusting the degree of a correspondence by trivial correspondences is achieved by correspondences coming from the range of the restriction map as in §5.6. A subtle failure of Fubini's theorem is what makes it possible to use these correspondences to adjust the degree in our case, as shown in §7.4. We then show in §7.5 that, in the case of function fields, one can recover the algebraic points of the curve inside the adèle class space as classical points of the periodic orbits of the dual system. This leads us to refine in §7.5.1 the notion described in §4.4 of passing to the dual system as an analog in characteristic zero of the unramified extension $\mathbb{K} \otimes_{\mathbb{F}_q} \bar{\mathbb{F}}_q$. In fact, we now show that there is a natural way to associate to a noncommutative space X not only its set of "classical points", but in a more subtle manner the set of all its classical points that can be defined over an unramified extension. This is obtained through the following basic steps:

$$X \overset{\text{Dual System}}{\longrightarrow} \hat{X} \overset{\text{Periodic Orbits}}{\longrightarrow} \cup \hat{X}_v \overset{\text{Classical Points}}{\longrightarrow} \cup \Xi_v = \Xi.$$

We also describe in §7.5.2, in the case of $\mathbb{K} = \mathbb{Q}$, the quantum statistical mechanical systems associated to the different valuations and how these classical points of the periodic orbits appear as low temperature KMS states for these systems. We describe the cyclic covering of the set of classical points of the periodic orbits in §7.5.3 and in §7.5.4 we obtain a global Morita equivalence that gives the valuation systems globally from the groupoid of the adèle class space (in fact a groupoid that differs from it only in the part that belongs to $C_{\mathbb{K}}$ which disappears upon "taking the complement" via the reduction map). We describe then in §7.5.5 an arithmetic subalgebra for the locus of classical points of the periodic orbits in terms of operators that behave like the Frobenius and local monodromy of the Weil–Deligne group at arithmetic infinity.

We then use in §7.6 the geometry of vortex equations and moduli spaces to justify thinking of the step in the Weil proof that makes a correspondence effective by modifying it via linear equations using the Riemann–Roch theorem in terms of achieving transversality (surjectivity) for a morphism of C^*-modules via a compact perturbation. We then assemble our tentative dictionary between the Weil proof and noncommutative geometry in §7.7.

We conclude this last chapter of the book by drawing in §8 a broad analogy between our approach to the Riemann zeta function and L-functions via the noncommutative geometry of $\mathbb{Q}$-lattices and its thermodynamical properties and the question of a good setup for quantum gravity in physics. Our starting point is a comparison in §8.1 between the role of spontaneous symmetry breaking in the origin of masses from the electroweak phase transition in physics and the phase transitions that occur in our general framework of interaction between quantum statistical mechanics and number theory described in Chapter 3. We propose in §8.2 and §8.4 the existence of a noncommutative algebra of coordinates on a space of “spectral correspondences” between product geometries of the type used in Chapter 1 in modelling particle physics, and a corresponding time evolution giving rise to KMS states that yield, at low temperature, the usual notion of geometry. This possible approach shows that gravity described by classical (pseudo)Riemannian geometry may be a low temperature phenomenon while geometry may disappear entirely in the high temperature regime, just as in the case of 2-dimensional $\mathbb{Q}$-lattices there are no KMS states above a certain temperature. This last section is more speculative in nature and we hope it will open the way to some future developments of the material collected in this book.

1. Morphisms and categories of noncommutative spaces

Noncommutative spaces are usually described in terms of some class of noncommutative algebras, although descriptions of noncommutative spaces are also possible via categorical or operadic constructions. It however soon becomes clear that morphisms of algebras (for instance $*$-homomorphisms of C^*-algebras) are inadequate as a notion of morphisms for a category of noncommutative spaces. This appears very clearly when one considers specific examples, such as the noncommutative tori, for which Morita equivalences play the role of isomorphisms. This suggests that a good notion of morphisms for noncommutative spaces should be based on correspondences, which include the algebra homomorphisms as a particular case, but which also accommodate more general type of mappings, such as those obtained by tensoring with a bimodule. We describe here two constructions available in the theory, which provide a way to embed a non-additive category of algebras in a pseudo-abelian category of noncommutative spaces. The first construction is Kasparov's KK-theory (cf. [**178**]) and the second is the abelian category of cyclic modules (cf. [**65**]).

1.1. The KK-category.

In this setting, we describe noncommutative spaces using separable C^*-algebras. The category of C^*-algebras with $*$-homomorphisms is not an additive category. However, one can introduce an *additive* category $\mathcal{KK}$, whose objects are separable C^*-algebras and with the morphisms given as follows. For $\mathcal{A}$ and $\mathcal{B}$ in $\mathrm{Obj}(\mathcal{KK})$, one has

$$\mathrm{Hom}(\mathcal{A},\mathcal{B}) = KK(\mathcal{A},\mathcal{B}), \tag{4.1}$$

where $KK(\mathcal{A},\mathcal{B})$ is Kasparov's bivariant K-theory ([**178**], cf. also §8 and §9.22 of [**20**]).

We describe in more detail the way the morphisms (4.1) are obtained.

First recall that a Hilbert module $\mathcal{E}$ over a C^*-algebra $\mathcal{B}$ is a right $\mathcal{B}$-module endowed with a positive $\mathcal{B}$-valued inner product with

$$\langle x,x\rangle \geq 0\,, \quad \langle xa, yb\rangle = a^*\langle x,y\rangle b, \quad \forall x,y\in\mathcal{E}, \quad \forall a,b\in\mathcal{B}, \tag{4.2}$$

with respect to which $\mathcal{E}$ is complete for the norm given by $||x|| = \sqrt{||\langle x,x\rangle||}$.

One can then introduce Kasparov modules: they are “Morita type” correspondences that generalize $*$-homomorphisms of C^*-algebras.

DEFINITION 4.1. *Suppose given a pair $\mathcal{A}$, $\mathcal{B}$ of separable C^*-algebras. A (bounded) Kasparov module is a triple $(\mathcal{E}, \phi, F)$, where*

- *$\mathcal{E}$ is a countably generated Hilbert module over $\mathcal{B}$.*
- *ϕ is a $*$-homomorphism, $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{E})$, with $\mathcal{B}(\mathcal{E})$ the algebra of bounded linear operators on $\mathcal{E}$. This gives $\mathcal{E}$ the $\mathcal{A}$–$\mathcal{B}$ bimodule structure.*
- *F is a bounded linear operator on $\mathcal{E}$ with the properties that the operators $[F, \phi(a)]$, $(F^2 - 1)\phi(a)$, and $(F^* - F)\phi(a)$ are in $\mathcal{K}(\mathcal{E})$, the ideal of compact operators, for all $a \in \mathcal{A}$.*

In the even case one further requires that

- $\mathcal{E}$ is $\mathbb{Z}/2$-graded, with grading operator $\gamma \in \mathcal{B}(\mathcal{E})$ with $\gamma^2 = 1$.
- $[\phi(a), \gamma] = 0$ for all $a \in \mathcal{A}$.
- F is odd i.e. $F\gamma = -\gamma F$.

One can form direct sums of Kasparov modules

$$(\mathcal{E}_0, \phi_0, F_0) \oplus (\mathcal{E}_1, \phi_1, F_1) = (\mathcal{E}_0 \oplus \mathcal{E}_1, \phi_0 \oplus \phi_1, F_0 \oplus F_1). \tag{4.3}$$

with $\gamma = \gamma_0 \oplus \gamma_1$ in the even case.

Given separable C^*-algebras $\mathcal{A}$ and $\mathcal{B}$, consider the C^*-algebra

$$I\mathcal{B} = \{f : [0,1] \to \mathcal{B} \,|\, f \text{ continuous}\},$$

with the endpoints evaluations $f_i : I\mathcal{B} \to \mathcal{B}$. Then on the set of Kasparov modules one introduces the homotopy equivalence relation in the following way.

DEFINITION 4.2. *Two Kasparov modules*

$$(\mathcal{E}_0, \phi_0, F_0) \sim_h (\mathcal{E}_1, \phi_1, F_1)$$

are homotopy equivalent if there exists an element $(\mathcal{E}, \phi, F)$, which is a Kasparov module for the C^*-algebras $\mathcal{A}$ and $I\mathcal{B}$, such that the element $(\mathcal{E} \otimes_{f_i} \mathcal{B}, f_i \circ \phi, f_i(F))$ is unitarily equivalent to $(\mathcal{E}_i, \phi_i, F_i)$. That is, there exists an element in the space of bounded operators from $\mathcal{E} \otimes_{f_i} \mathcal{B}$ to $\mathcal{E}_i$ which is a unitary intertwining the morphisms $f_i \circ \phi$ and ϕ_i and the operators $f_i(F)$ and F_i.

The KK-groups are then defined as follows.

DEFINITION 4.3. *$KK(\mathcal{A}, \mathcal{B})$ is the abelian group obtained as the Grothendieck group of the semigroup of even Kasparov modules up to (even) homotopy with the operation of direct sum* (4.3)*.*

One usually assumes that the C^*-algebra $\mathcal{B}$ is σ-unital, to avoid the possible problem of not having enough countably generated Hilbert modules. Moreover, in the σ-unital case, the different equivalence relations one can consider on the set of Kasparov modules all coincide (cf. e.g. [**20**] §17.2 and §17.3).

The element of $KK(\mathcal{A}, \mathcal{B})$ determined by a $*$-homomorphism of C^*-algebras $\phi : \mathcal{A} \to \mathcal{B}$ is simply given by

$$(\mathcal{B}, \phi, F = 0), \qquad \text{with} \quad \gamma = 1.$$

It also has a representative with $F^2 = 1$ of the form ([**20**], Example 17.1.2)

$$(\mathcal{B} \oplus \mathcal{B}, \phi \oplus 0, F), \qquad \text{with} \quad F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; \gamma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The composition of correspondences defined by Kasparov modules is given by the cup product

$$\otimes_{\mathcal{D}} \; : \; KK(\mathcal{A},\mathcal{D}) \otimes KK(\mathcal{D},\mathcal{B}) \to KK(\mathcal{A},\mathcal{B})\,. \tag{4.4}$$

One also has external tensor products

$$KK(\mathcal{A},\mathcal{B}) \otimes KK(\mathcal{C},\mathcal{D}) \to KK(\mathcal{A}\otimes\mathcal{C},\mathcal{B}\otimes\mathcal{D}). \tag{4.5}$$

These admit a simpler description in terms of *unbounded* Kasparov modules (cf. [**12**]). One modifies Definition 4.1 in the following way.

DEFINITION 4.4. *Unbounded Kasparov modules are triples $(\mathcal{E},\phi,D)$, with $(\mathcal{E},\phi)$ as in Definition 4.1, where D is an* unbounded *regular self-adjoint operator D on $\mathcal{E}$ such that*

- *$(1+D^2)^{1/2}\phi(a)$ extends to an element of $\mathcal{K}(\mathcal{E})$, for all $a\in\mathcal{A}$.*
- *$[D,\phi(a)]$ is densely defined and extends to an element of $\mathcal{B}(\mathcal{E})$ for all a in a dense subalgebra of $\mathcal{A}$.*

The operator $D(1+D^2)^{-1/2}$ extends to a bounded operator F such that $(\mathcal{E},\phi,F)$ defines a bounded Kasparov module as in Definition 4.1 (cf. [**20**], Prop. 17.11.3).

Then the external tensor product (4.4) is obtained ([**12**], cf. [**68**] §IV, Appendix A) by taking $\mathcal{E}_1\otimes\mathcal{E}_2$ and $D_1\otimes 1+\gamma\otimes D_2$ in the even case with $\mathbb{Z}/2\mathbb{Z}$-grading γ and in the odd case

$$D = \begin{pmatrix} 0 & D_1\otimes 1 - i\otimes D_2 \\ D_1\otimes 1 + i\otimes D_2 & 0 \end{pmatrix}.$$

REMARK 4.5. *The category $\mathcal{KK}$ whose objects are separable C^*-algebras and whose morphisms are given by* (4.1) *is an additive category.*

In fact, the category $\mathcal{KK}$ obtained in this way is a universal enveloping additive category for the category of C^*-algebras (cf. [**20**] §9.22.1). One can obtain a pseudo-abelian category, through the following general procedure.

Given an additive category $\mathcal{C}$, the pseudo-abelian envelope (Karoubian envelope) $\Psi\mathcal{C}$ is the category whose objects are pairs (A,e), with $A\in\mathrm{Obj}(\mathcal{C})$ and an idempotent $e=e^2\in\mathrm{End}(A)=\mathrm{Hom}_{\mathcal{C}}(A,A)$. The morphisms are of the form

$$\mathrm{Hom}_{\Psi\mathcal{C}}((A,e),(A',e')) = e'\mathrm{Hom}_{\mathcal{C}}(A,A')e.$$

The functor $\mathcal{C}\to\Psi\mathcal{C}$ that sends $A\mapsto(A,1)$ is fully faithful.

Moreover, a recent result of Meyer and Nest [**225**] shows that the additive category $\mathcal{KK}$ is in fact a triangulated category. The shift is obtained from the suspension $\mathcal{A}\mapsto\Sigma\mathcal{A}=C_0(\mathbb{R})\otimes\mathcal{A}$. This requires introducing more general objects of the form $(\mathcal{A},n)$ with $\mathcal{A}\in\mathrm{Obj}(\mathcal{KK})$ and $n\in\mathbb{Z}$ with

$$\mathrm{Hom}((\mathcal{A},n),(\mathcal{B},m)) = \varinjlim_{k\in\mathbb{N}} \mathrm{Hom}_{\mathcal{KK}}(\Sigma^{n+k}\mathcal{A},\Sigma^{m+k}\mathcal{B}).$$

The shift then is of the form $(\mathcal{A},n)\mapsto(\mathcal{A},n+1)$. Given a $*$-homomorphism $\phi:\mathcal{A}\to\mathcal{B}$, the mapping cone $C(\phi)$ is defined as

$$C(\phi) := \{(a,f)\in\mathcal{A}\times C_0((0,1),\mathcal{B})\,|\,\phi(a)=f(1)\}, \tag{4.6}$$

and fits in a diagram

$$\Sigma\mathcal{A} \xrightarrow{\iota} C(\phi) \xrightarrow{\pi} \mathcal{A} \xrightarrow{\phi} \mathcal{B} \tag{4.7}$$

of $*$-homomorphisms. The distinguished triangles are diagrams of the form

$$\Sigma\mathcal{B}' \to \mathcal{C}' \to \mathcal{A}' \to \mathcal{B}'$$

that are isomorphic to a cone diagram (4.7), namely

$$\begin{tikzcd} \Sigma\mathcal{B} \arrow[r, "\iota"] \arrow[d, "\Sigma f_B", "\cong"'] & C(\phi) \arrow[r, "\pi"] \arrow[d, "f_C", "\cong"'] & \mathcal{A} \arrow[r] \arrow[d, "f_A", "\cong"'] & \mathcal{B} \arrow[d, "f_B", "\cong"'] \\ \Sigma\mathcal{B}' \arrow[r] & \mathcal{C}' \arrow[r] & \mathcal{A}' \arrow[r] & \mathcal{B}'. \end{tikzcd} \tag{4.8}$$

It is proved in [**225**] that the category obtained in this way is a triangulated category. The result in fact is proved there more generally in the equivariant case.

We next discuss a way of embedding the non-additive category of noncommutative algebras in an abelian category, where one does not work with C^*-algebras and functional analysis, but more generally with noncommutative algebras over a field $\mathbb{K}$. This construction is based on cyclic modules (cf. [**65**]). We first introduce the cyclic category (cf. [**68**], §III, Appendix A).

1.2. The cyclic category.

The cyclic category Λ is a refinement of the simplicial category (cf. [**65**]). Namely, one has one object $[n] \in \mathrm{Obj}(\Lambda)$ for each non-negative integer and the morphisms are generated by face and degeneracy morphisms $\delta_i : [n-1] \to [n]$ and $\sigma_j : [n+1] \to [n]$ with relations

$$\delta_j\delta_i = \delta_i\delta_{j-1} \quad \text{for } i < j, \qquad \sigma_j\sigma_i = \sigma_i\sigma_{j+1}, \quad i \le j \tag{4.9}$$

$$\sigma_j\delta_i = \begin{cases} \delta_i\sigma_{j-1} & i < j \\ 1_n & \text{if } i = j \text{ or } i = j+1 \\ \delta_{i-1}\sigma_j & i > j+1, \end{cases} \tag{4.10}$$

as in the case of the simplicial category, with additional morphisms $\tau_n : [n] \to [n]$ with relations

$$\begin{aligned} &\tau_n\delta_i = \delta_{i-1}\tau_{n-1} \quad 1 \le i \le n, \quad \tau_n\delta_0 = \delta_n \\ &\tau_n\sigma_i = \sigma_{i-1}\tau_{n+1} \quad 1 \le i \le n, \quad \tau_n\sigma_0 = \sigma_n\tau_{n+1}^2 \\ &\tau_n^{n+1} = 1_n. \end{aligned} \tag{4.11}$$

Remark 4.6. *The classifying space of the cyclic category Λ satisfies*

$$B\Lambda = BU(1) = \mathbb{CP}^\infty, \tag{4.12}$$

This result (cf. Theorem 10 of [**65**] and Proposition 7.2.7 of [**211**]) should be seen in contrast with the case of the simplicial category, whose nerve is a contractible space. Thus, the cyclic category can be thought of as a "discretization of the circle" and cyclic cohomology is closely related to S^1-equivariant cohomology.

The cyclic category has another equivalent description in terms of its cyclic covering. One defines a category $\tilde{\Lambda}$ with objects of the form $(\mathbb{Z}, n) \in \mathrm{Obj}(\tilde{\Lambda})$. The morphisms

$(\mathbb{Z}, n) \to (\mathbb{Z}, m)$ are given by nondecreasing maps $f : \mathbb{Z} \to \mathbb{Z}$ satisfying the periodicity condition $f(x+n) = f(x) + m$. There is an action of $\mathbb{Z}$ on $\tilde{\Lambda}$ by translations and the quotient $\Lambda = \tilde{\Lambda}/\mathbb{Z}$ gives back the cyclic category described above.

DEFINITION 4.7. *Given a category $\mathcal{C}$, a cyclic object in $\mathcal{C}$ is a contravariant functor $\Lambda \to \mathcal{C}$. A cyclic object in the category* $\mathrm{Vect}_{\mathbb{K}}$ *is called a $\mathbb{K}(\Lambda)$-module or a cyclic module.*

PROPOSITION 4.8. *The category of $\mathbb{K}(\Lambda)$-modules whose objects are the cyclic modules, i.e. the contravariant functors $\Lambda \to \mathrm{Vect}_{\mathbb{K}}$ and whose morphisms are the natural transformations is an abelian category.*

PROOF. One can in fact construct a ring $\mathbb{K}(\Lambda)$ generated over $\mathbb{K}$ by idempotents e_n, $e_n e_m = 0$ for $n \neq m$, and symbols $\sigma_j^n, \delta_i^n, \tau_n$ fulfilling relations (4.9), (4.10) and (4.11), with e_n instead of 1_n. One requires that $\delta_i^n = \delta_i^n e_{n-1} = e_n \delta_i^n$, $\sigma_j^n = \sigma_j^n e_{n+1} = e_n \sigma_j^n$ while $\tau_n = \tau_n e_n = e_n \tau_n$. A contravariant functor $\Lambda \to \mathrm{Vect}_{\mathbb{K}}$ is then equivalently described as a right $\mathbb{K}(\Lambda)$-module. Since the modules over a ring form an abelian category one obtains the desired result. ☐

The reason why this construction is useful in noncommutative geometry lies in the fact that algebras define corresponding cyclic modules, as the following result of [**65**] shows.

LEMMA 4.9. *A unital algebra $\mathcal{A}$ over a field $\mathbb{K}$ determines a $\mathbb{K}(\Lambda)$-module $\mathcal{A}^\natural$.*

PROOF. The contravariant functor $\mathcal{A}^\natural : \Lambda \to \mathrm{Vect}_{\mathbb{K}}$ is defined on objects by

$$\mathcal{A}^\natural : [n] \mapsto \mathcal{A}^{\otimes(n+1)} = \mathcal{A} \otimes \mathcal{A} \otimes \cdots \otimes \mathcal{A} \tag{4.13}$$

and on morphisms by

$$\mathcal{A}^\natural : \delta_i \mapsto (a^0 \otimes \cdots \otimes a^n) \mapsto (a^0 \otimes \cdots \otimes a^i a^{i+1} \otimes \cdots \otimes a^n) \tag{4.14}$$

$$\mathcal{A}^\natural : \sigma_j \mapsto (a^0 \otimes \cdots \otimes a^n) \mapsto (a^0 \otimes \cdots \otimes a^j \otimes 1 \otimes a^{j+1} \otimes \cdots \otimes a^n) \tag{4.15}$$

$$\mathcal{A}^\natural : \tau_n \mapsto (a^0 \otimes \cdots \otimes a^n) \mapsto (a^n \otimes a^0 \otimes \cdots \otimes a^{n-1}). \tag{4.16}$$

☐

A morphism $\phi : \mathcal{A} \to \mathcal{B}$ of unital $\mathbb{K}$-algebras determines a morphism $\phi^\natural : \mathcal{A}^\natural \to \mathcal{B}^\natural$ of the corresponding cyclic modules by

$$\phi^\natural(a^0 \otimes \cdots \otimes a^n) = \phi(a^0) \otimes \cdots \otimes \phi(a^n)\,.$$

One obtains in this way an embedding of the (non-additive) category of unital algebras over $\mathbb{K}$ into the abelian category of $\mathbb{K}(\Lambda)$-modules.

When passing from algebras to the corresponding $\mathbb{K}(\Lambda)$-modules, one obtains many more morphisms than those induced by algebra homomorphisms. An important example of morphisms of cyclic modules that are not induced by morphisms of algebras are traces. Suppose given a trace

$$\tau : \mathcal{A} \to \mathbb{K} \tag{4.17}$$

on a unital algebra $\mathcal{A}$ over $\mathbb{K}$. This is a linear functional with $\tau(1) = 1$ satisfying $\tau(ab) = \tau(ba)$. In the case of topological algebras, one requires τ to be continous. Such a functional defines a corresponding morphism of cyclic modules

$$\tau^\natural : \mathcal{A}^\natural \to \mathbb{K}^\natural, \tag{4.18}$$

of the form

$$\tau^\natural(x^0 \otimes \cdots \otimes x^n) = \tau(x^0 \cdots x^n). \tag{4.19}$$

In fact one has more generally,

Proposition 4.10. *Given unital algebras $\mathcal{A}, \mathcal{B}$, a trace on $\mathcal{A}$ determines a morphism of cyclic modules*

$$\tau^\natural = \mathrm{Tr} : (\mathcal{B} \otimes \mathcal{A})^\natural \to \mathcal{B}^\natural \tag{4.20}$$

$$\mathrm{Tr}((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)) = x_0 \otimes \cdots \otimes x_n \, \tau(t_0 \cdots t_n). \tag{4.21}$$

Proof. For $0 \le i \le n-1$ one has

$$\begin{aligned}\mathrm{Tr}\,\delta_i((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)) &= x_0 \otimes \cdots \otimes x_i x_{i+1} \otimes \cdots \otimes x_n \, \tau(t_0\, t_1 \, \cdots \, t_n) \\ &= \delta_i \,\mathrm{Tr}((x_0 \otimes t_0) \otimes (x_1 \otimes t_1) \otimes \cdots \otimes (x_n \otimes t_n)).\end{aligned}$$

Similarly,

$$\begin{aligned}&\mathrm{Tr}\,\delta_n((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)) = x_n\, x_0 \otimes \cdots \otimes x_{n-1}\, \tau(t_n\, t_0\, t_1 \cdots t_{n-1}) \\ &= \delta_n(x_0 \otimes \cdots \otimes x_n)\, \tau(t_0\, t_1 \, \cdots \, t_n) = \delta_n \,\mathrm{Tr}((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)).\end{aligned}$$

In the same way, for $0 \le j \le n$, one has

$$\begin{aligned}&\mathrm{Tr}\,\sigma_j((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)) = x_0 \otimes \cdots \otimes x_j \otimes 1 \otimes x_{j+1} \otimes \cdots \otimes x_n\, \tau(t_0\, t_1 \cdots t_n) \\ &= \sigma_j(x_0 \otimes \cdots \otimes x_n)\, \tau(t_0\, t_1 \cdots t_n) = \sigma_j \,\mathrm{Tr}((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)).\end{aligned}$$

Finally, one has

$$\begin{aligned}\mathrm{Tr}\,\tau_n((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)) &= x_n \otimes x_0 \otimes \cdots \otimes x_{n-1}\, \tau(t_0\, t_1 \cdots t_n) \\ &= \tau_n \,\mathrm{Tr}((x_0 \otimes t_0) \otimes \cdots \otimes (x_n \otimes t_n)).\end{aligned}$$

∎

Another important example of morphisms in the category of cyclic modules comes from correspondences given by bimodules.

Corollary 4.11. *Let $\mathcal{E}$ be an $\mathcal{A}$–$\mathcal{B}$ bimodule, which is finite projective as a right $\mathcal{B}$-module. Then there is an associated cyclic morphism*

$$\mathcal{E}^\natural \in \mathrm{Hom}(\mathcal{A}^\natural, \mathcal{B}^\natural).$$

Proof. Since $\mathcal{E}$ is finite projective it is a direct summand of the right module $\mathcal{B}^n$ for some n, so that $\mathrm{End}_{\mathcal{B}}(\mathcal{E}) \subset M_n(\mathcal{B}) = \mathcal{B} \otimes M_n(\mathbb{C})$. Thus one obtains $\tau^\natural$ from the left action of $\mathcal{A}$, which gives a homomorphism ρ from $\mathcal{A}$ to $\mathrm{End}_{\mathcal{B}}(\mathcal{E})$ and hence to $\mathcal{B} \otimes M_n(\mathbb{C})$. One then takes the composition of $\rho^\natural$ with the trace map from $\mathcal{B} \otimes M_n(\mathbb{C})$ to $\mathcal{B}$ which is also a cyclic morphism as follows from Proposition 4.10 applied to the usual trace on $M_n(\mathbb{C})$. The obtained cyclic morphism is of the form

$$\mathcal{E}^\natural = \mathrm{Tr}^\natural \circ \rho^\natural : \mathcal{A}^\natural \to \mathcal{B}^\natural. \tag{4.22}$$

∎

REMARK 4.12. *Note that the above cyclic morphism depends upon the choice of a connection on the right $\mathcal{B}$-module $\mathcal{E}$, which is used in order to write $\mathcal{E}$ as a direct summand of the right module $\mathcal{B}^n$ for some n.*

1.3. The non-unital case.

The construction of cyclic modules associated to unital algebras extends to the non-unital case. In fact, if we think of a non-unital associative algebra as a non-compact noncommutative space, one can consider various possible compactifications and then the associated cyclic modules. This means considering a unital algebra $\mathcal{A}^{\rm comp}$ with the property that the non-unital algebra $\mathcal{A}$ is contained in it as an essential ideal $\mathcal{A} \subset \mathcal{A}^{\rm comp}$ (cf. Chapter 3, Definition 3.2).
A particular case is the "one-point compactification" $\mathcal{A}^{\rm comp} = \tilde{\mathcal{A}}$ obtained by adjoining a unit to $\mathcal{A}$.
To an inclusion $\mathcal{A} \subset \mathcal{A}^{\rm comp}$ as above, one can associate a $\mathbb{K}(\Lambda)$-module $(\mathcal{A}, \mathcal{A}^{\rm comp})^\natural$ defined as follows. To an object $[n]$ of Λ it associates the subspace of the $\mathbb{K}$-vector space $(\mathcal{A}^{\rm comp})^{\otimes(n+1)}$ of elements of the form

$$\sum a_0 \otimes \cdots \otimes a_n$$

where at least one of the entries $a_j \in \mathcal{A}^{\rm comp}$ belongs to $\mathcal{A}$.
In the case of the one-point compactification this means elements

$$\sum a_0 \otimes \cdots \otimes a_n, \quad \text{with} \ \ a_j \in \tilde{\mathcal{A}}, \quad \text{such that} \quad \sum \varepsilon(a_0) \cdots \varepsilon(a_n) = 0,$$

where $\varepsilon : \tilde{\mathcal{A}} \to \mathbb{C}$ is the augmentation with $\operatorname{Ker} \varepsilon = \mathcal{A}$.
The images of the morphisms δ_n, σ_n, τ_n are defined as in (4.14), (4.15), (4.16).
For instance, one can consider as a non-unital algebra the ideal $\mathcal{L}^1$ of trace class operators on a Hilbert space (here $\mathbb{K} = \mathbb{C}$). This example will be useful later in §4.9 below.

1.4. Cyclic (co)homology.

The Ext functors (the right derived functors of Hom) in the abelian category of $\mathbb{K}(\Lambda)$-modules, when applied to cyclic modules associated to algebras, give the cyclic cohomology, by the following result of [**65**].

PROPOSITION 4.13. *Let $\mathcal{A}$ be a unital associative algebra over a field $\mathbb{K}$ and $\mathcal{A}^\natural$ the corresponding $\mathbb{K}(\Lambda)$-module. Then the cyclic cohomology of the algebra $\mathcal{A}$ is given by*

$$HC^n(\mathcal{A}) = \operatorname{Ext}^n_{\mathbb{K}(\Lambda)}(\mathcal{A}^\natural, \mathbb{K}^\natural). \tag{4.23}$$

The identification with $\operatorname{Ext}^n(\mathcal{A}^\natural, \mathbb{K}^\natural)$ follows by constructing a double complex of projective $\mathbb{K}(\Lambda)$-modules that is a resolution of the trivial $\mathbb{K}(\Lambda)$-module $\mathbb{K}^\natural$ (cf. [**65**]).
Similarly, one obtains cyclic homology in terms of Tor functors (the left derived functors of $\otimes$), namely

$$HC_n(\mathcal{A}) = \operatorname{Tor}_n(\mathbb{K}^\natural, \mathcal{A}^\natural). \tag{4.24}$$

See [**67**], [**65**], [**211**] for more details. This yields the following bicomplex whose total homology computes the cyclic homology for a unital algebra $\mathcal{A}$. Let $\mathcal{A}^\natural([n]) =$

$\mathcal{A}^{\otimes(n+1)}$ as above with $\mathcal{A}^\natural_n(\delta_i) : \mathcal{A}^\natural([n]) \to \mathcal{A}^\natural([n-1])$, $\mathcal{A}^\natural_n(\sigma_i) : \mathcal{A}^\natural([n]) \to \mathcal{A}^\natural([n+1])$ and $\mathcal{A}^\natural_n(\tau_n) : \mathcal{A}^\natural([n]) \to \mathcal{A}^\natural([n])$ the corresponding morphisms as in (4.14), (4.15), (4.16). The bicomplex whose total homology computes the cyclic homology is then given by

$$\begin{tikzcd}
\vdots \arrow[d, "b_{n+1}"] & \vdots \arrow[d, "b_{n+1}"] & \vdots \arrow[d, "b_{n+1}"] & \\
\mathcal{A}^\natural([n]) \arrow[r, "1-\lambda_n"] \arrow[d, "-b'_n"] & \mathcal{A}^\natural([n]) \arrow[r, "N_n"] \arrow[d, "-b'_n"] & \mathcal{A}^\natural([n]) \arrow[r, "1-\lambda_n"] \arrow[d, "b_n"] & \cdots \\
\mathcal{A}^\natural([n-1]) \arrow[r, "1-\lambda_{n-1}"] \arrow[d, "b_{n-1}"] & \mathcal{A}^\natural([n-1]) \arrow[r, "N_{n-1}"] \arrow[d, "b_{n-1}"] & \mathcal{A}^\natural([n-1]) \arrow[r, "1-\lambda_{n-1}"] \arrow[d, "b_{n-1}"] & \cdots \\
\vdots & \vdots & \vdots &
\end{tikzcd}$$

where we define $b_n, b'_n : \mathcal{A}^\natural([n]) \to \mathcal{A}^\natural([n-1])$ by

$$b_n = \sum_{i=0}^{n} (-1)^i \mathcal{A}^\natural_n(\delta_i), \quad b'_n = \sum_{i=0}^{n-1} (-1)^i \mathcal{A}^\natural_n(\delta_i), \tag{4.25}$$

and we set $\lambda_n = (-1)^n \mathcal{A}^\natural_n(\tau_n)$ and

$$N_n = \sum_{i=0}^{n} \lambda_n^i. \tag{4.26}$$

The horizontal and vertical differentials d_1, d_2 defined in this way satisfy $d_1^2 = d_2^2 = 0$ and $d_1 d_2 = -d_2 d_1$. This bicomplex is closely related to the (b, B) bicomplex of Chapter 1 §10.4.

Notice that, passing to the dual spaces by assigning to each integer $n \geq 0$ the vector space C^n of $(n+1)$-linear forms $\varphi(x^0, \ldots, x^n)$ on $\mathcal{A}$, with the basic operations $\delta_i : C^{n-1} \to C^n$ and $\sigma_i : C^{n+1} \to C^n$ given by

$$\begin{aligned}
(\delta_i \varphi)(x^0, \ldots, x^n) &= \varphi(x^0, \ldots, x^i x^{i+1}, \ldots, x^n), \ i = 0, 1, \ldots, n-1 \\
(\delta_n \varphi)(x^0, \ldots, x^n) &= \varphi(x^n x^0, x^1, \ldots, x^{n-1}) \\
(\sigma_j \varphi)(x^0, \ldots, x^n) &= \varphi(x^0, \ldots, x^j, 1, x^{j+1}, \ldots, x^n), \ j = 0, 1, \ldots, n \\
(\tau_n \varphi)(x^0, \ldots, x^n) &= \varphi(x^n, x^0, \ldots, x^{n-1}).
\end{aligned} \tag{4.27}$$

one obtains a covariant functor $\mathcal{A}^\sharp : \Lambda \to \mathrm{Vect}_{\mathbb{K}}$. This, using the canonical isomorphism $\Lambda \simeq \Lambda^{op}$ of [**65**], can be seen as a contravariant functor, i.e. a cyclic module. One obtains in this way a similar description of cyclic cohomology as

$$HC^n(\mathcal{A}) = \mathrm{Ext}^n(\mathbb{K}^\sharp, \mathcal{A}^\sharp). \tag{4.28}$$

The formula (4.28) for cyclic cohomology can be compared with the proposed *absolute cohomology* for motives that we discussed in our brief overview of motives in

Chapter 1. The absolute cohomology is computed also by an Ext functor

$$H^n_{\mathrm{abs}}(M) = \mathrm{Ext}^n_{\mathcal{D}(\mathbb{K})}(1, M) \tag{4.29}$$

in a triangulated category $\mathcal{D}(\mathbb{K})$ of mixed motives over the field $\mathbb{K}$. The actual category of mixed motives would be obtained as the heart of a t-structure in this triangulated category. Notice how, in the case of cyclic cohomology, the Ext is instead taken directly in the abelian category of $\mathbb{K}(\Lambda)$-modules. We like to think of cyclic (co)homology as a kind of absolute cohomology for noncommutative spaces.

Remark 4.14. All of the discussion above, regarding the cyclic module associated to an algebra, extends to the context of locally convex topological algebras. In that context one needs to work with topological tensor products in the construction of $\mathcal{A}^\natural$ and of continuous multilinear forms in the construction of $\mathcal{A}^\sharp$.
Moreover, an important issue arises, since the ranges of continuous linear maps are not necessarily closed subspaces. In order to preserve the duality between cyclic homology and cyclic cohomology we shall define the cokernel of a cyclic map $T : \mathcal{A}^\natural \to \mathcal{B}^\natural$ as the quotient of $\mathcal{B}^\natural$ by the closure of the range of T. In a dual manner, the kernel of the transposed map $T^t : \mathcal{B}^\sharp \to \mathcal{A}^\sharp$ is automatically closed and is the dual of the above.

2. Endomotives

As we discussed already in the overview on motives in Chapter 1, the case of Artin motives presents the simplest possible construction of a category of pure motives. It is sufficiently interesting though to recover the absolute Galois group as its motivic Galois group, as we already saw in Chapter 1 §8. We recall here to begin with a few well known facts about Artin motives (cf. e.g. [**2**], [**282**]).

In this chapter we work with motives over a field $\mathbb{K}$, which we assume to be a *number field*, i.e. a finite algebraic extension of $\mathbb{Q}$. In the following we fix an embedding $\sigma : \mathbb{K} \hookrightarrow \mathbb{C}$ and, for a variety X over $\mathbb{K}$, we write $X(\mathbb{C})$ for $\sigma X(\mathbb{C})$. We take $\bar{\mathbb{K}}$ to be the algebraic closure of $\sigma(\mathbb{K}) \subset \mathbb{C}$ in $\mathbb{C}$.

Let $\mathcal{V}^0_{\mathbb{K}}$ denote the category of zero-dimensional smooth projective varieties, i.e. finite dimensional reduced commutative algebras A over $\mathbb{K}$. This means that $X = \mathrm{Spec}(A)$ consists of a finite set of points. For $X \in \mathrm{Obj}(\mathcal{V}^0_{\mathbb{K}})$ we denote by $X(\bar{\mathbb{K}})$ the set of algebraic points of X.
One has

$$X(\bar{\mathbb{K}}) = \mathrm{Hom}_{\mathrm{Alg}_{\mathbb{K}}}(A, \bar{\mathbb{K}}). \tag{4.30}$$

In fact, in general for any unital abelian $\mathbb{K}$-algebra B the set of B-points of X is the set of homomorphisms of unital abelian $\mathbb{K}$-algebras

$$X(B) = \mathrm{Hom}_{\mathrm{Alg}_{\mathbb{K}}}(A, B). \tag{4.31}$$

The set $X(\bar{\mathbb{K}})$ as in (4.30) is a finite set on which $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$ acts continuously by $\chi \in X(\bar{\mathbb{K}}) \mapsto g \circ \chi$ for all $g \in G$.
Consider the map

$$\mathrm{Obj}(\mathcal{V}^0_{\mathbb{K}}) \ni X \mapsto X(\bar{\mathbb{K}}) \in \mathrm{Obj}(\mathcal{S}_G) \tag{4.32}$$

that assigns to a zero-dimensional variety its set of algebraic points. The Grothendieck–Galois correspondence states that (4.32) is a functor that gives an equivalence of categories between the category $\mathcal{V}^0_{\mathbb{K}}$ and the category $\mathcal{S}_G$ of finite sets with a continuous action of $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$.

The category $\mathcal{AM}_{\mathbb{K},\mathbb{E}}$ of Artin motives over $\mathbb{K}$ with coefficients in a field $\mathbb{E}$ of characteristic zero is the subcategory of the category $\mathcal{M}_{\mathrm{num},\mathbb{E}}(\mathbb{K})$ of pure motives with numerical equivalence and with coefficients in $\mathbb{E}$ generated by the objects of $\mathcal{V}^0_{\mathbb{K}}$. We simply write $\mathcal{AM}_{\mathbb{K}}$ for $\mathcal{AM}_{\mathbb{K},\mathbb{Q}}$.

More explicitly, given X and Y in $\mathrm{Obj}(\mathcal{V}^0_{\mathbb{K}})$, one defines $M(X,Y)$ to be the finite dimensional $\mathbb{Q}$-vector space

$$M(X,Y) := \mathrm{Hom}_G(\mathbb{Q}^{X(\bar{\mathbb{K}})}, \mathbb{Q}^{Y(\bar{\mathbb{K}})}) = (\mathbb{Q}^{X(\bar{\mathbb{K}})\times Y(\bar{\mathbb{K}})})^G. \tag{4.33}$$

An algebraic cycle, or a (virtual) correspondence, is therefore given by an element $Z \in M(X,Y)$, i.e. a formal linear combination of connected components of $X \times Y$ with coefficients in $\mathbb{Q}$. These are identified with $\mathbb{Q}$-valued G-invariant functions on $X(\bar{\mathbb{K}}) \times Y(\bar{\mathbb{K}})$ by $Z = \sum a_i Z_i \mapsto \sum a_i \mathbf{1}_{Z_i(\bar{\mathbb{K}})}$, where $\mathbf{1}_{Z_i(\bar{\mathbb{K}})}$ is the characteristic function, and the Z_i are connected components of $X \times Y$.

In general, an element $U \in M(X,Y)$ is of the form $U = \sum_i a_i Z_i$, where the Z_i are components in the product $X \times Y$. The composition of correspondences is $\mathbb{Q}$-bilinear and given as usual by the intersection product on the Z_i. Suppose given components $Z \subset X_1 \times X_2$ and $Z' \subset X_2 \times X_3$. The composition

$$\circ : M(X_1, X_2) \times M(X_2, X_3) \to M(X_1, X_3)$$

is given by

$$Z \circ Z' = \pi_{13*}(\pi_{12}^* Z \bullet \pi_{23}^* Z') \tag{4.34}$$

with $\bullet$ the intersection product and $\pi_{ij} : X_1 \times X_2 \times X_3 \to X_i \times X_j$ the projections.

In the zero-dimensional case assigning such an element $U \in M(X,Y)$ is equivalent to the datum of the corresponding map $U_* \in \mathrm{Hom}_G(\mathbb{Q}^{X(\bar{\mathbb{K}})}, \mathbb{Q}^{Y(\bar{\mathbb{K}})})$ and the composition (4.34) corresponds to the composition of the maps

$$(U \circ V)_* = V_* \circ U_*. \tag{4.35}$$

The category of Artin motives is the (pseudo)abelian envelope of the additive category with objects $\mathrm{Obj}(\mathcal{V}^0_{\mathbb{K}})$ and morphisms $M(X,Y) \otimes_{\mathbb{Q}} \mathbb{E}$. It is in fact an abelian category, since $\mathcal{M}_{\mathrm{num},\mathbb{E}}(\mathbb{K})$ is abelian. The choice of the equivalence relation is irrelevant in the case of Artin motives. In fact, in dimension zero there is no equivalence relation needed on the cycles in the product. The category $\mathcal{AM}_{\mathbb{K}}$ is a neutral Tannakian category with fiber functor

$$\omega : X \mapsto H^0(X(\mathbb{C}), \mathbb{Q}) = \mathbb{Q}^{X(\bar{\mathbb{K}})}. \tag{4.36}$$

The fiber functor gives an equivalence of categories of $\mathcal{AM}_{\mathbb{K}}$ with the category Rep_G of finite dimensional $\mathbb{Q}$-linear representations of the absolute Galois group $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$. (Here $\mathbb{K}$ is of characteristic zero.) This can be seen as a "linearized version" of the Grothendieck–Galois correspondence mentioned above.

The classical realizations of Artin motives are given by

$$H_B(\sigma X) = \mathbb{Q}^{X(\bar{\mathbb{K}})}, \quad H_{\mathrm{et}}(\bar{X}) = \mathbb{Q}_\ell^{X(\bar{\mathbb{K}})} \quad \text{and} \quad H_{\mathrm{dR}}(X) = A,$$

for $X = \mathrm{Spec}(A)$, with A a finite dimensional reduced $\mathbb{K}$-algebra.

REMARK 4.15. In passing from $\mathcal{V}^0_{\mathbb{K}}$ to the category $\mathcal{AM}_{\mathbb{K}}$ one in general needs to add new objects, to obtain the (pseudo)abelian envelope. One can see this in a very simple example, for $\mathbb{K} = \mathbb{Q}$. Consider the field $\mathbb{L} = \mathbb{Q}(\sqrt{2})$. Then consider the one-dimensional nontrivial representation of the Galois group $G = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ that factors through the character of order two of $\mathrm{Gal}(\mathbb{L}/\mathbb{Q})$. This representation does not correspond to an object in $\mathrm{Obj}(\mathcal{V}^0_{\mathbb{Q}})$ but can be obtained as the range of the projector $p = (1-\sigma)/2$, where σ is the generator of $\mathrm{Gal}(\mathbb{L}/\mathbb{Q})$. Namely, this is a new object in $\mathrm{Obj}(\mathcal{AM}_{\mathbb{Q}})$ which is not an object of $\mathcal{V}^0_{\mathbb{Q}}$.

An observation that will be useful in the following is that, if we view the objects in the category of Artin motives as reduced finite dimensional commutative algebras, then the correspondences defined above in terms of algebraic cycles in the product can be given an equivalent description in terms of bimodules.

LEMMA 4.16. *For X and Y in $\mathrm{Obj}(\mathcal{V}^0_{\mathbb{K}})$, an element $U \in M(X,Y)$ defines a (virtual) bimodule $\mathcal{E}_U$, compatibly with the composition.*

PROOF. We have $X = \mathrm{Spec}\, A$, $Y = \mathrm{Spec}\, B$, for A and B reduced finite dimensional algebras over $\mathbb{K}$. We can write the element $U \in M(X,Y)$ in the form $U = \sum_i a_i Z_i$ with $a_i \in \mathbb{Q}$ and Z_i components in $X \times Y$. One obtains a virtual bimodule $\mathcal{E}_U = \sum_i a_i \mathcal{E}_{Z_i}$ where $\mathcal{E}_{Z_i}$ is the actual A–B bimodule obtained from $Z_i \subset X \times Y$ using the restrictions of the projection maps to get the bimodule structure on C_i, $Z_i = \mathrm{Spec}\, C_i$. More precisely, with $f = p_X|_{Z_i}$ $g = p_Y|_{Z_i}$, the restrictions, and $\tilde{f}$, $\tilde{g}$ the associated morphisms of algebras, the bimodule structure of $C = C_i$ is given by

$$a\,\xi\, b = \tilde{f}(a)\,\xi\,\tilde{g}(b)\,, \quad \forall a \in A\,,\ b \in B\,,\ \xi \in C\,. \tag{4.37}$$

By (4.35) it is clear that the composition (4.34) corresponds to the tensor product of bimodules

$$\mathcal{E}_Z \otimes_B \mathcal{E}_{Z'} \simeq \mathcal{E}_{Z \circ Z'}. \tag{4.38}$$

□

In the following we construct a faithful embedding of the category $\mathcal{AM}_{\mathbb{K},\mathbb{E}}$ of Artin motives in an additive category of noncommutative spaces whose objects are given by projective systems of Artin motives with the action of semigroups of endomorphisms. We call such objects *endomotives*. We give two versions of the category of endomotives, one purely algebraic and defined over any field $\mathbb{K}$, and another one which is analytic and uses the C^*-algebra description of noncommutative spaces. We show that, for $\mathbb{K}$ a number field, the category of algebraic endomotives embeds in the category of analytic endomotives.

2.1. Algebraic endomotives.

Let X_α be a projective system of zero-dimensional varieties, that is, a collection $\{X_\alpha\}_{\alpha\in I}$ of objects in $\mathcal{V}^0_{\mathbb{K}}$, with a projective system of morphisms $\xi_{\alpha,\beta} : X_\beta \to X_\alpha$, and a countable partially ordered indexing set I. We denote by X the projective limit

$$X = \varprojlim_{\alpha} X_\alpha, \tag{4.39}$$

with maps $\xi_\alpha : X \to X_\alpha$. We view X as a pro-variety. Equivalently, we consider a directed system of algebras A_α over $\mathbb{K}$, with $X_\alpha = \operatorname{Spec} A_\alpha$. The direct limit

$$A = \varinjlim_\alpha A_\alpha \tag{4.40}$$

is a commutative unital algebra over $\mathbb{K}$, no longer finite dimensional.
The graphs $\Gamma(\xi_{\alpha,\beta})$ of the morphisms of the projective system are G-invariant subsets of $X_\beta(\bar{\mathbb{K}}) \times X_\alpha(\bar{\mathbb{K}})$. They define a compatible system of morphisms in $M(X_\alpha, X_\beta)$, with respect to composition (4.34).

Consider then a unital abelian semigroup S acting on A by algebra endomorphisms $\rho : A \to A$. Notice that the endomorphisms in general need not preserve the levels A_α. Also we do not assume that algebra endomorphisms send the unit to itself. In fact, in general the image $\rho(1) = e$ is an idempotent in the algebra, $e^2 = e$.
We assume that the endomorphisms $\rho \in S$ induce isomorphisms $A \xrightarrow{\simeq} eAe$, where the reduced algebra eAe is here the same as the range eA of the projection since A is commutative. This is equivalent to the condition that the elements $\rho \in S$ induce isomorphisms $x \in X^e \mapsto x \cdot \rho \in X$, where $X^e \subset X$ is the subvariety associated to the idempotent $e = \rho(1)$, i.e. corresponding to the compressed algebra eAe. The notation $x \cdot \rho$ for this action of S on X is convenient when viewing the points $x \in X(\bar{\mathbb{K}})$ as characters of the algebra A, since the action on points is then simply given by composition of the characters with the endomorphisms. For $\rho, \rho_j \in S$ the equality $\rho\rho_1 = \rho\rho_2$ implies that $\rho_1 = \rho_2$ since ρ is an injective algebra homomorphism. Thus we can embed S as a sub-semigroup of the enveloping abelian group $\tilde{S}$.

Definition 4.17. *Let (X_α, S) be the data of a projective system of Artin motives and a unital abelian semigroup S with the properties described above. An algebraic endomotive is a unital associative algebra over $\mathbb{K}$ of the form*

$$\mathcal{A}_{\mathbb{K}} = A \rtimes S, \tag{4.41}$$

where A is the abelian $\mathbb{K}$-algebra (4.40).

The crossed product algebra (4.41) has an explicit presentation obtained by adjoining to A new generators U_ρ and U_ρ^*, for $\rho \in S$, satisfying the relations

$$\begin{aligned} &U_\rho^* U_\rho = 1, && U_\rho U_\rho^* = \rho(1), && \forall \rho \in S \\ &U_{\rho_1 \rho_2} = U_{\rho_1} U_{\rho_2}, && U^*_{\rho_2 \rho_1} = U^*_{\rho_1} U^*_{\rho_2}, && \forall \rho_1, \rho_2 \in S \\ &U_\rho\, a = \rho(a)\, U_\rho, && a\, U_\rho^* = U_\rho^*\, \rho(a), && \forall \rho \in S, \forall a \in A. \end{aligned} \tag{4.42}$$

Lemma 4.18. (1) *The algebra* (4.41) *is the linear span of the monomials $U^*_{\rho_1} a U_{\rho_2}$, for $a \in A$ and $\rho_j \in S$.*
(2) *The product $U_s = U^*_{\rho_2} U_{\rho_1}$ only depends upon the ratio $s = \rho_1/\rho_2 \in \tilde{S}$.*
(3) *The algebra* (4.41) *is the linear span of the monomials $a\, U_s$, $s \in \tilde{S}$, $a \in A$.*

Proof. (1) In fact, it is enough to show that we have

$$U^*_{\rho_1} a U_{\rho_2} U^*_{\rho_3} b U_{\rho_4} = U^*_{\rho_1\rho_3} \rho_3(a) \rho_2\rho_3(1) \rho_2(b) U_{\rho_2\rho_4}, \ \forall \rho_j \in S\,, a, b \in A\,. \tag{4.43}$$

which follows from (4.42) since the semigroup S is abelian. To see this one first replaces $\rho_2\rho_3(1)$ by the product $U_{\rho_2\rho_3} U^*_{\rho_2\rho_3}$. One then has

$$U^*_{\rho_1\rho_3} \rho_3(a) U_{\rho_2\rho_3} = U^*_{\rho_1} a U_{\rho_2}$$

using $U_{\rho_3} a = \rho_3(a) U_{\rho_3}$ and $U^*_{\rho_3} U_{\rho_3} = 1$. Applying the same procedure to the other terms yields (4.43).
(2) For $\rho \in S$ one has

$$U^*_{\rho\rho_2} U_{\rho\rho_1} = U^*_{\rho_2} U^*_\rho U_\rho U_{\rho_1} = U^*_{\rho_2} U_{\rho_1} \,.$$

This suffices to show (2).
(3) One has $U^*_\rho = U^*_\rho U_\rho U^*_\rho = U^*_\rho \rho(1)$. Thus since ρ is an isomorphism with the reduced algebra eA, for $e = \rho(1)$, we can write any product $U^*_\rho a = U^*_\rho e a = b U^*_\rho$ with $b = \rho^{-1}(ae) \in A$. □

For $s = \rho_1/\rho_2 \in \tilde{S}$ the projections

$$E(s) = \rho_1^{-1}(\rho_2(1)\rho_1(1)) \quad \text{and} \quad F(s) = \rho_2^{-1}(\rho_2(1)\rho_1(1)) \tag{4.44}$$

only depend on s and the map $(\rho_2)^{-1}\rho_1$ is an isomorphism of the reduced algebras

$$s = \rho_2^{-1}\rho_1 \,:\, A_{E(s)} \to A_{F(s)} \,. \tag{4.45}$$

For idempotents e, f in a commutative algebra we write $e \leq f$ when $ef = e$. One has $e \leq f_1 f_2$ if and only if $e \leq f_j$ for all j.

Lemma 4.19. *Let $E(s)$ and $F(s)$ be the projections of* (4.44). *They satisfy*

$$E(s^{-1}) = F(s) = s(E(s)) \,, \quad F(s\,s') \geq F(s) s(F(s')) \,, \quad \forall s, s' \in \tilde{S} \,. \tag{4.46}$$

Proof. To prove the last inequality, let $s = \rho_1/\rho_2$, $s' = \rho'_1/\rho'_2$. Since $\rho_2\rho'_2$ is injective it is enough to show that

$$\rho_2\rho'_2(F(s)s(F(s'))) \leq \rho_2\rho'_2(F(ss')) = (\rho_2\rho'_2(1))(\rho_1\rho'_1(1))$$

One has

$$\rho_2(F(s)s(F(s'))) = \rho_2(1)\rho_1(1)\rho_1(F(s')) = \rho_2(1)\rho_1(F(s'))$$

and the conclusion follows from

$$\rho'_2\rho_1(F(s')) = \rho_1\rho'_2(F(s')) = \rho_1(\rho'_2(1)\rho'_1(1)) \leq \rho_1\rho'_1(1) \,.$$

□

Moreover, the partial isomorphism s^{-1} is the inverse of s, and one has

$$U_s\, a = s(a)\, U_s \,, \quad \forall a \in A_{E(s)} \,. \tag{4.47}$$

By Lemma 4.18 the following map is a linear isomorphism:

$$a = (a_s) \in \oplus_{s\in\tilde{S}} A_{F(s)} \mapsto \sum a_s\, U_s \in \mathcal{A}_{\mathbb{K}} \,. \tag{4.48}$$

The linear isomorphism (4.48) allows one to interpret the product in $\mathcal{A}_{\mathbb{K}}$ as the convolution product for a groupoid $\mathcal{G}$ which is constructed as follows.

Definition 4.20. *As a space $\mathcal{G}$ is the disjoint union*

$$\mathcal{G} = \sqcup_{s\in\tilde{S}} X^{F(s)} \tag{4.49}$$

which corresponds to the algebra $\oplus_{s\in\tilde{S}} A_{F(s)}$, which is the commutative algebra direct sum of the reduced algebras $A_{F(s)}$. The range and source maps are given by the natural projection from $\mathcal{G}$ to X and by its composition with the antipode $\mathbf{S}$ *which is given, at the algebra level, by*

$$\mathbf{S}(a)_s = s(a_{s^{-1}}) \,, \quad \forall s \in \tilde{S} \,. \tag{4.50}$$

The composition in the groupoid corresponds to the product of monomials

$$a\,U_s\,b\,U_t = \,a\,s(b)\,U_{st}\,. \tag{4.51}$$

Thus, an algebraic endomotive $\mathcal{A}_{\mathbb{K}}$ can be equivalently described through the data of the associated groupoid $\mathcal{G} = \mathcal{G}(X_\alpha, S)$. This is useful in order to define the notion of correspondence for endomotives.

As we saw above the natural definition of correspondences in the context of motives [**2**] (cf. §8 of Chapter 1) is as algebraic cycles, that is, combinations of subvarieties in the product. In the context of KK-theory, one also has the notion of geometric correspondences as in [**98**] (cf. §3.1 below) obtained from smooth manifolds that project to the product. We shall now extend this notion to the framework of endomotives. In that framework one is not just dealing with spaces but one needs to also account for the presence of group actions. This is best done using groupoids because they generalize both the "space part" of the correspondence (the cycle) and the group action. Thus given algebraic endomotives (X_α, S) and $(X'_{\alpha'}, S')$, with associated groupoids $\mathcal{G} = \mathcal{G}(X_\alpha, S)$, $\mathcal{G}' = \mathcal{G}(X'_\alpha, S')$ a geometric correspondence is given by a $\mathcal{G}$–$\mathcal{G}'$-space Z where the right action of $\mathcal{G}'$ fulfills a suitable *étale* condition which we now describe.

Given a space such as $\mathcal{G}$, i.e. a disjoint union $Z = \operatorname{Spec} C$ of zero-dimensional pro-varieties over $\mathbb{K}$, a *right action* of $\mathcal{G}$ on Z is specified by a map $g : Z \to X$ and a collection of partial isomorphisms

$$z \in g^{-1}(F(s)) \mapsto z \cdot s \in g^{-1}(E(s)) \tag{4.52}$$

fulfilling the obvious rules for a partial action of the abelian group $\tilde{S}$. More precisely one requires that

$$g(z \cdot s) = g(z) \cdot s\,, \quad z \cdot (ss') = (z \cdot s) \cdot s' \text{ on } g^{-1}(F(s) \cap s(F(s'))) \tag{4.53}$$

where we denote as above by $x \mapsto x \cdot s$ the given partial action of $\tilde{S}$ on X.

One checks that such an action gives to the $\mathbb{K}$-linear space C (for $Z = \operatorname{Spec} C$) a structure of right module over $\mathcal{A}_{\mathbb{K}}$.

DEFINITION 4.21. *We say that the action of $\mathcal{G}$ on Z is* étale *if the corresponding module C is finite and projective over $\mathcal{A}_{\mathbb{K}}$.*

This notion immediately extends to left actions and we use the obvious terminology of $\mathcal{G}$–$\mathcal{G}'$ space Z for a pair of commuting left and right actions on Z of the groupoid $\mathcal{G}$.

DEFINITION 4.22. *Given algebraic endomotives (X_α, S) and $(X'_{\alpha'}, S')$, an étale correspondence is a $\mathcal{G}(X_\alpha, S)$–$\mathcal{G}(X'_{\alpha'}, S')$-space Z such that the right action of $\mathcal{G}(X'_{\alpha'}, S')$ is étale. We define the $\mathbb{Q}$-linear space of (virtual) correspondences $M((X_\alpha, S), (X'_{\alpha'}, S'))$ as formal linear combinations $U = \sum_i a_i Z_i$ of étale correspondences Z_i modulo isomorphism and the equivalence $Z \sqcup Z' \sim Z + Z'$ where $Z \sqcup Z'$ is the disjoint union.*

The composition of correspondences is given by the fiber product over a groupoid. Namely, for algebraic endomotives (X_α, S), $(X'_{\alpha'}, S')$, $(X''_{\alpha''}, S'')$ and correspondences Z and W, their composition is given by

$$Z \circ W = Z \times_{\mathcal{G}'} W, \tag{4.54}$$

which is the fiber product over the groupoid $\mathcal{G}' = \mathcal{G}(X'_{\alpha'}, S')$. We can then form a category of algebraic endomotives in the following way.

DEFINITION 4.23. *The category* $\mathcal{EV}^0_{\mathbb{K},\mathbb{E}}$ *of algebraic endomotives is the pseudo-abelian envelope of the additive category whose objects are algebraic endomotives* (X_α, S) *and with morphisms*

$$\mathrm{Hom}_{\mathcal{EV}^0_{\mathbb{K},\mathbb{E}}}((X_\alpha,S),(X'_{\alpha'},S')) = M((X_\alpha,S),(X'_{\alpha'},S'))\otimes_{\mathbb{Q}}\mathbb{E}. \tag{4.55}$$

It is a tensor category with the tensor product induced by

$$(X_\alpha,S)\otimes(X'_{\alpha'},S') = (X_\alpha\times X'_{\alpha'}, S\times S'). \tag{4.56}$$

2.2. Analytic endomotives.

When we consider as above a projective system X_α and the corresponding directed system of algebras A_α, we can assign to the pro-variety $X = \varprojlim X_\alpha$ its set of algebraic points $X(\bar{\mathbb{K}})$. This is a totally disconnected compact Hausdorff space

$$\mathcal{X} = X(\bar{\mathbb{K}}) = \varprojlim_\alpha X_\alpha(\bar{\mathbb{K}}) \tag{4.57}$$

in the topology of the projective limit. Thus, by the usual Gelfand–Naimark correspondence, the space $\mathcal{X}$ is equivalently described by the abelian C^*-algebra $C(\mathcal{X})$.

If (X_α, S) are the data of an algebraic endomotive, then the semigroup S acts by homeomorphisms $\chi\in\mathcal{X}^e\mapsto\chi\circ\rho\in\mathcal{X}$, where $\mathcal{X}^e = X^e(\bar{\mathbb{K}})$ and $e=\rho(1)$. Thus, S acts on the C^*-algebra $C(\mathcal{X})$ by the endomorphisms

$$\rho(f)(\chi) = 0 \text{ if } \chi(e)=0\,,\quad \rho(f)(\chi) = f(\chi\circ\rho) \text{ if } \chi(e)=1\,,$$

and one can consider the semigroup crossed product C^*-algebra

$$C(\mathcal{X})\rtimes S. \tag{4.58}$$

See [**195**] and [**192**] and the references therein for the general theory of crossed products by semigroups.

This crossed product can be described equivalently as the groupoid C^*-algebra $C^*(\mathcal{G}(\bar{\mathbb{K}}))$ of the locally compact étale groupoid $\mathcal{G}(\bar{\mathbb{K}})$ corresponding to the $\bar{\mathbb{K}}$-points of the groupoid $\mathcal{G}(X_\alpha,S)$ of (4.49). An element γ of $\mathcal{G}(\bar{\mathbb{K}})$ is given by a pair (χ,s) where $s\in\tilde{S}$ and χ is a character of the reduced algebra $A_{F(s)}$ i.e. a character of A such that $\chi(F(s))=1$. The range and source maps are given by

$$\mathrm{range}(\chi,s)=\chi\,,\quad \mathrm{source}(\chi,s)=\chi\circ s \tag{4.59}$$

and both are elements of $X(\bar{\mathbb{K}})$. The composition $\gamma_1\circ\gamma_2$ makes sense when the source $\chi_1\circ s_1$ of γ_1 is the range χ_2 of γ_2 and it is then given by

$$\gamma_1\circ\gamma_2 = (\chi_1, s_1s_2)\,. \tag{4.60}$$

One has to check that $\chi_1(F(s_1s_2))=1$, but $F(s_1s_2)\geq F(s_1)s_1(F(s_2))$ by (4.46) and $\chi_1\circ s_1=\chi_2$ which gives the required equality. One checks that one obtains in this way a locally compact étale groupoid $\mathcal{G}(\bar{\mathbb{K}})$.

Using the chosen embedding $\sigma:\mathbb{K}\hookrightarrow\mathbb{C}$, we can embed the algebra $A=\varinjlim A_\alpha$ as a subalgebra $A\subset C(X)$:

$$A\hookrightarrow C(\mathcal{X}),\quad a\mapsto f_a,\quad f_a(\chi)=\chi(a), \tag{4.61}$$

where we write $f_a\in C(\mathcal{X})$ for the locally constant function corresponding to the element $a\in A$. This embedding extends to the algebra $A_{\mathbb{C}} = A\otimes_{\mathbb{K}}\mathbb{C}$. Notice that

A is not in general an involutive subalgebra for the canonical involution of $C(\mathcal{X})$. It is true, however, that the $\mathbb{C}$-linear span $A_{\mathbb{C}} \subset C(\mathcal{X})$ is an involutive subalgebra of $C(\mathcal{X})$ since it is the algebra of locally constant functions on $\mathcal{X}$. This can be seen since the $\mathbb{C}$-algebra $A_\alpha \otimes_{\mathbb{K}} \mathbb{C}$ is the sum of finitely many copies of $\mathbb{C}$ indexed by the characters of A_α with values in $\bar{\mathbb{K}}$. The C^*-algebra $C(\mathcal{X})$ is an abelian AF algebra which is the C^*-completion of the direct limit $A_{\mathbb{C}} = \varinjlim A_\alpha \otimes_{\mathbb{K}} \mathbb{C}$ of finite dimensional abelian algebras.

We obtain a corresponding subalgebra of the crossed product C^*-algebra

$$\mathcal{A}_{\mathbb{C}} = \mathcal{A}_{\mathbb{K}} \otimes \mathbb{C} = A_{\mathbb{C}} \rtimes S \subset C(\mathcal{X}) \rtimes S. \tag{4.62}$$

We say in this case that the algebra $\mathcal{A}_{\mathbb{K}} = A \rtimes S$ over $\mathbb{K}$ (the algebraic endomotive) gives an *arithmetic structure* on the analytic endomotive described by the C^*-algebra $C(\mathcal{X}) \rtimes S$.

DEFINITION 4.24. *Suppose given the data $(\mathcal{X}, S)$ of a totally disconnected compact Hausdorff space $\mathcal{X}$ and a unital abelian semigroup S acting by homeomorphisms $\mathcal{X}^{\mathrm{e}} \to \mathcal{X}$ as above. Then an analytic endomotive is the étale locally compact groupoid $\mathcal{G}(\mathcal{X}, S)$.*

To an analytic endomotive is associated the crossed product C^*-algebra

$$C^*(\mathcal{G}) = C(\mathcal{X}) \rtimes S. \tag{4.63}$$

Note that since the semigroup S (and the group $\tilde{S}$) is abelian there is no distinction between the reduced and maximal norms so that the crossed product C^*-algebra is unambiguously defined. The norm is obtained as

$$||f|| = \mathrm{Sup}_{\mathcal{X}}\, ||\pi_x(f)|| \tag{4.64}$$

where for each $x \in \mathcal{X}$ one lets π_x be the representation by left convolution in the Hilbert space $\ell^2(\mathcal{G}_x)$ of the countable fiber over $x \in \mathcal{X}$ of the source map.

We construct correspondences of analytic endomotives in the same way as for algebraic ones but in the locally compact context. Given a totally disconnected locally compact space $\mathcal{Z}$ we define a right action of $\mathcal{G} = \mathcal{G}(\mathcal{X}, S)$ on $\mathcal{Z}$ as a continuous map g from $\mathcal{Z}$ to $\mathcal{X}$ and a collection indexed by $s \in \tilde{S}$ of partial isomorphisms as in (4.52),

$$z \in g^{-1}(F(s)) \mapsto z \circ s \in g^{-1}(E(s)) \tag{4.65}$$

fulfilling the obvious rules for a partial action of the abelian group $\tilde{S}$. More precisely one requires that

$$g(z \circ s) = g(z) \circ s\,, \quad z \circ (ss') = (z \circ s) \circ s' \text{ on } g^{-1}(F(s) \cap s(F(s'))) \tag{4.66}$$

where we denote as above by $x \mapsto x \circ s$ the given partial action of $\tilde{S}$ on $\mathcal{X}$.

A right action of $\mathcal{G}$ on $\mathcal{Z}$ gives on the space $C_c(\mathcal{Z})$ of continuous functions with compact support on $\mathcal{Z}$ a structure of right module over $C_c(\mathcal{G})$ as in the algebraic case. When the fibers of the map g are discrete (countable) subsets of $\mathcal{Z}$ one can define on $C_c(Z)$ an inner product with values in $C_c(\mathcal{G})$ by

$$\langle \xi, \eta \rangle (x, s) = \sum_{z \in g^{-1}(x)} \bar{\xi}(z)\, \eta(z \circ s) \tag{4.67}$$

We define the notion of *étale* action by

Definition 4.25. *A right action of $\mathcal{G}$ on $\mathcal{Z}$ is* étale *if and only if the fibers of the map g are discrete and the identity is a compact operator in the right C^*-module $\mathcal{E}_{\mathcal{Z}}$ over $C^*(\mathcal{G})$ given by* (4.67).

This will be automatic if the action comes from an algebraic étale action since in that case the right C^*-module $\mathcal{E}_{\mathcal{Z}}$ over $C^*(\mathcal{G})$ is finite and projective since it is induced by the finite projective module over $\mathcal{A}_{\mathbb{K}}$ of Definition 4.21.
Exactly as in the algebraic context we now obtain the notion of étale correspondence.

Definition 4.26. *Given analytic endomotives $(\mathcal{X}, S)$ and $(\mathcal{X}', S')$, an étale correspondence is a $\mathcal{G}$–$\mathcal{G}'$-space $\mathcal{Z}$ such that the right action of $\mathcal{G}'$ is étale. The $\mathbb{Q}$-linear space of (virtual) correspondences $M((\mathcal{X}, S), (\mathcal{X}', S'))$ is given by formal linear combinations $U = \sum_i a_i \mathcal{Z}_i$ of étale correspondences $\mathcal{Z}_i$ modulo isomorphism and the equivalence $\mathcal{Z} \sqcup \mathcal{Z}' \sim \mathcal{Z} + \mathcal{Z}'$, where $\mathcal{Z} \sqcup \mathcal{Z}'$ is the disjoint union.*

The composition of correspondences is given by the fiber product over a groupoid. Namely, for

$$U \in M((\mathcal{X}, S), (\mathcal{X}', S')) \quad \text{and} \quad V \in M((\mathcal{X}', S'), (\mathcal{X}'', S''))$$

with $U = \sum a_i \mathcal{Z}_i$ and $V = \sum b_j \mathcal{W}_j$ we have

$$\mathcal{Z}_i \circ \mathcal{W}_j = \mathcal{Z}_i \times_{\mathcal{G}'} \mathcal{W}_j, \tag{4.68}$$

the fiber product over the groupoid $\mathcal{G}' = \mathcal{G}(\mathcal{X}', S')$.
The category of analytic endomotives is then obtained in the following way.

Definition 4.27. *The category $C^*\mathcal{V}^0_{\mathbb{K},\mathbb{E}}$ is the (pseudo)abelian envelope of the additive category with objects the analytic endomotives $(\mathcal{X}, S)$ and morphisms*

$$\mathrm{Hom}_{C^*\mathcal{V}^0_{\mathbb{K},\mathbb{E}}}((\mathcal{X}, S), (\mathcal{X}', S')) = M((\mathcal{X}, S), (\mathcal{X}', S')) \otimes_{\mathbb{Q}} \mathbb{E}. \tag{4.69}$$

An important property of analytic endomotives is that, even though we are working with C^*-algebras, if the space $\mathcal{X} = X(\bar{\mathbb{K}})$ comes from an underlying algebraic endomotive, then the analytic object inherits a Galois action of $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$, as we show in the next subsection.

2.3. Galois action.

Suppose given, as above, a projective system X_α of zero-dimensional smooth projective varieties over a number field $\mathbb{K}$. Consider the pro-variety $X = \varprojlim X_\alpha$. The set of algebraic points $X(\bar{\mathbb{K}})$ has the equivalent description in terms of characters of the algebra $A = \varinjlim A_\alpha$, as in (4.30) and (4.31),

$$X(\bar{\mathbb{K}}) = \{\chi : A \to \bar{\mathbb{K}}\}. \tag{4.70}$$

Thus, we obtain an action of the Galois group $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$ on $X(\bar{\mathbb{K}})$ by composition, namely for $g \in G$ and $\chi \in X(\bar{\mathbb{K}})$ we define $g(\chi) = g \circ \chi$ as the composite

$$A \xrightarrow{\chi} \bar{\mathbb{K}} \xrightarrow{g} \bar{\mathbb{K}}. \tag{4.71}$$

The following result then shows that we have a Galois action on analytic endomotives.

LEMMA 4.28. *The action* (4.71) *of* $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$ *on* $X(\bar{\mathbb{K}})$ *defines a canonical action of* G *by automorphisms of the* C^**-algebra* $C(\mathcal{X}) \rtimes S$*, preserving globally the abelian subalgebra* $C(\mathcal{X})$ *and fixing the* U_ρ *and* U_ρ^*.

PROOF. The action (4.71) gives an action of G by automorphisms of the algebra $C(\mathcal{X})$ in the form

$$g : f \mapsto f \circ g^{-1}, \quad \forall g \in G, \quad \forall f \in C(\mathcal{X}). \tag{4.72}$$

The action (4.71) commutes with the action of the semigroup S since the latter acts by precomposition on A. First, for $\rho \in S$, the idempotent $e = \rho(1)$ belongs to A and for $g \in G$ and $\chi \in X(\bar{\mathbb{K}})$ we have

$$g(\chi)(e) = g \circ \chi(e) = \chi(e)$$

since $\chi(e) \in \{0, 1\}$. Thus $e = \rho(1)$ is G-invariant. Moreover, we have

$$A \xrightarrow{\rho} A \xrightarrow{\chi} \bar{\mathbb{K}} \xrightarrow{g} \bar{\mathbb{K}}, \tag{4.73}$$

for all $g \in G$ and all $\rho \in S$, and for all $\chi \in X(\bar{\mathbb{K}}) = \mathcal{X}$, so that we have

$$(g \circ \chi) \circ \rho = g \circ (\chi \circ \rho)$$

and the two actions commute. □

We obtain the following result on the intertwining of states and Galois action.

PROPOSITION 4.29. *Consider algebras* $\mathcal{A}_{\mathbb{K}} = A \rtimes S$ *and* $C(\mathcal{X}) \rtimes S$ *as above. Then for any pure state* φ *on the abelian* C^**-algebra* $C(\mathcal{X})$ *the values* $\varphi(A) \subset \bar{\mathbb{K}}$ *and the Galois action satisfies*

$$g\,\varphi(a) = \varphi(g^{-1}(a)), \quad \forall a \in A, \quad \forall g \in G. \tag{4.74}$$

PROOF. An element $\chi \in X(\bar{\mathbb{K}}) = \mathrm{Hom}_{\mathrm{Alg}_{\mathbb{K}}}(A, \bar{\mathbb{K}})$ determines a pure state φ_χ on the abelian C^*-algebra $C(\mathcal{X})$ defined by evaluation

$$\varphi_\chi(f) := f(\chi)\,, \tag{4.75}$$

and all pure states of $C(\mathcal{X})$ are of this form. The Galois action on $X(\bar{\mathbb{K}})$ gives a corresponding action on states by

$$g\varphi(f) = \varphi(g^{-1}(f)), \tag{4.76}$$

where the action on $C(\mathcal{X})$ is of the form (4.72).
For $a \in A$ and f_a the corresponding locally constant function in $C(\mathcal{X})$ we have

$$\varphi_\chi(f_a) = f_a(\chi) = \chi(a) \in \bar{\mathbb{K}} \tag{4.77}$$

and the Galois action is then of the form

$$g\varphi_\chi(f_a) = g \circ \chi(a) = g(\chi)(a) = f_a(g(\chi)) = \varphi_\chi(g^{-1} f_a). \tag{4.78}$$

□

In the case of normal abelian extensions we have the following stronger result on the intertwining of states and Galois actions.

PROPOSITION 4.30. *Suppose given $\mathcal{A}_{\mathbb{K}} = A \rtimes S$ and $C(\mathcal{X}) \rtimes S$ as above, with $A = \varinjlim A_\alpha$. Assume that the algebras A_α are of the form $A_\alpha = \prod_i \mathbb{L}_{\alpha,i}$ where the $\mathbb{L}_{\alpha,i}$ are abelian and normal field extensions of $\mathbb{K}$. Then the subalgebras A and $\mathcal{A}_{\mathbb{K}}$ are globally invariant under the action of G on the C^*-algebra $C(\mathcal{X}) \rtimes S$ and the states φ on $C(\mathcal{X}) \rtimes S$ induced by pure states of $C(\mathcal{X})$ fulfill the intertwining equation*

$$g\,\varphi(a) = \varphi(\theta(g)(a)), \quad \forall a \in \mathcal{A}_{\mathbb{K}}, \quad \forall g \in G^{\mathrm{ab}} = G/[G,G]\,, \tag{4.79}$$

where $\theta(g) = g^{-1}$.

PROOF. First notice that, by the assumption, the field $\mathbb{L}_{\alpha,i}$ is a normal and abelian extension of $\mathbb{K}$. Thus an automorphism of the field $\mathbb{L}_{\alpha,i}$ of the form

$$\tilde{g} = \chi^{-1} \circ g \circ \chi, \tag{4.80}$$

with $g \in G$, is independent of the choice of $\chi \in \mathrm{Hom}_{\mathrm{Alg}_{\mathbb{K}}}(\mathbb{L}_{\alpha,i}, \bar{\mathbb{K}})$. We can then define an automorphism $\tilde{g} \in \mathrm{Aut}(A_\alpha)$ such that

$$\chi \circ \tilde{g} = g \circ \chi\,, \quad \forall \chi \in \mathrm{Hom}_{\mathrm{Alg}_{\mathbb{K}}}(A_\alpha, \bar{\mathbb{K}}) \tag{4.81}$$

We then have

$$f_{\tilde{g}(a)}(\chi) = \chi(\tilde{g}(a)) = g(\chi(a)) = g^{-1}(f_a)(\chi) \tag{4.82}$$

which shows that $f_{\tilde{g}(a)} = g^{-1}(f_a)$. This shows that $A \subset C(\mathcal{X})$ is globally invariant under the action of G. Thus, the algebra $\mathcal{A}_{\mathbb{K}} \otimes \mathbb{C} \subset C(\mathcal{X}) \rtimes S$ is also globally invariant under G since the action of G fixes the U_ρ and U_ρ^*. Moreover, notice that $g \mapsto \theta(g) = g^{-1}$ induces a group homomorphism on the abelianization $G/[G,G]$. ∎

In the normal non-abelian case the global invariance of A no longer holds. One can define an action of G on A but only in a non-canonical manner which depends on the choices of base points.

2.4. Uniform systems and measured endomotives.

Another piece of data that we want to consider for analytic endomotives is a measure. This will define a state on the resulting algebra. It will be evident in §4 why we want to have the data of an algebra together with a state.

Given a probability measure space $(\mathcal{X}, \mu)$ and a map $f : \mathcal{X} \to \mathcal{Y}$, the probability measure $f_*\mu$ on $\mathcal{Y}$ is defined by $f_*\mu(A) = \mu(f^{-1}(A))$. Given a projective system $\mathcal{X}_\alpha$ with maps $\xi_{\alpha,\beta} : \mathcal{X}_\beta \to \mathcal{X}_\alpha$, the Prokhorov extension of measures shows that the compatibility condition

$$\mu_\alpha = (\xi_{\alpha,\beta})_* \mu_\beta \tag{4.83}$$

is necessary and sufficient for the probability measures μ_α on $\mathcal{X}_\alpha$ to determine a probability measure μ on the projective limit $\mathcal{X} = \varprojlim_\alpha \mathcal{X}_\alpha$ satisfying

$$(\xi_\alpha)_* \mu = \mu_\alpha, \tag{4.84}$$

for the projection maps $\xi_\alpha : \mathcal{X} \to \mathcal{X}_\alpha$. When condition (4.83) is satisfied we say that $(\mathcal{X}_\alpha, \mu_\alpha)$ is a projective system of probability measure spaces. Notice that the limit measure μ depends on the full projective system and not only on the projective limit $\mathcal{X}$.

DEFINITION 4.31. *Let X_α be a projective system of objects in $\mathcal{V}^0_{\mathbb{K}}$. Consider the finite sets $\mathcal{X}_\alpha = X_\alpha(\bar{\mathbb{K}})$. We say that the projective system is uniform if the normalized counting measures μ_α on $\mathcal{X}_\alpha$ form a projective system of measure spaces satisfying (4.83). We say that an algebraic endomotive (X_α, S) is uniform if the X_α form a uniform projective system.*

A uniform system X_α determines a probability measure μ on $\mathcal{X} = X(\bar{\mathbb{K}})$. Thus, we say that a uniform algebraic endomotive (X_α, S) determines a *measured* analytic endomotive $(\mathcal{X}, S, \mu)$, in the sense of the following definition.

DEFINITION 4.32. *A measured analytic endomotive $(\mathcal{X}, S, \mu)$ is an analytic endomotive $(\mathcal{X}, S)$ together with a probability measure μ on $\mathcal{X}$ with the property that for all $s \in S$ the Radon–Nikodým derivatives*

$$\frac{ds^*\mu}{d\mu} \tag{4.85}$$

are locally constant functions on $\mathcal{X}$.

To take into account the extra datum of the measure μ it is convenient to require that correspondences of measured analytic endomotives also have a compatibility condition with the measure.

DEFINITION 4.33. *Let $(\mathcal{X}, S, \mu)$ and $(\mathcal{X}', S', \mu')$ be two measured analytic endomotives. A correspondence in $M((\mathcal{X}, S, \mu), (\mathcal{X}', S', \mu'))$ is an element $U = \sum_i a_i \mathcal{Z}_i \in M((\mathcal{X}, S), (\mathcal{X}', S'))$ with the condition that under the left action of $s \in S$ on $\mathcal{Z}_i$ the Radon–Nikodým derivatives satisfy*

$$\frac{d(s^*\mu_{\mathcal{Z}_i})}{d\mu_{\mathcal{Z}_i}} = \frac{d(s^*\mu)}{d\mu}, \tag{4.86}$$

where $\mu_{\mathcal{Z}_i} = g^*_{\mathcal{X}'}\mu'$ is the pullback measure with the counting measure on the fibers.

Thus, one can consider categories of endomotives, where one restricts to the subcategories generated by uniform algebraic endomotives and by measured analytic endomotives with the correspondences as in Definition 4.33.

2.5. Compatibility of endomotives categories.

The following result relates the categories of analytic endomotives, algebraic endomotives, and Artin motives. It also describes how they map to the KK-category.

THEOREM 4.34. *Let $\mathcal{AM}_{\mathbb{K},\mathbb{E}}$, $\mathcal{EV}^0_{\mathbb{K},\mathbb{E}}$, $C^*\mathcal{V}^0_{\mathbb{E}}$ denote respectively the categories of Artin motives, algebraic endomotives, and analytic endomotives, as above.*

(1) *The map $\mathcal{G} \mapsto \mathcal{G}(\bar{\mathbb{K}})$ determines a tensor functor*

$$\mathcal{P} : \mathcal{EV}^0_{\mathbb{K},\mathbb{E}} \to C^*\mathcal{V}^0_{\mathbb{E}} \tag{4.87}$$

from algebraic to analytic endomotives.

(2) *The Galois group $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$ acts by natural transformations of the functor $\mathcal{P}$.*

(3) *The category $\mathcal{AM}_{\mathbb{K},\mathbb{E}}$ of Artin motives is a full subcategory of $\mathcal{EV}^0_{\mathbb{K},\mathbb{E}}$.*

(4) *The composite functor*

$$k \circ \mathcal{P} : \mathcal{EV}^0_{\mathbb{K},\mathbb{E}} \to \mathcal{KK} \otimes \mathbb{E} \tag{4.88}$$

maps the full subcategory $\mathcal{AM}_{\mathbb{K},\mathbb{E}}$ of Artin motives faithfully to the category $\mathcal{KK}_{G,\mathbb{E}}$ of G-equivariant KK-theory with coefficients in $\mathbb{E}$.

Proof. (1) The functor $\mathcal{P}$ defined by taking points over $\bar{\mathbb{K}}$ maps

$$\mathcal{P} : (X_\alpha, S) \mapsto (\mathcal{X}, S) \tag{4.89}$$

where $\mathcal{X} = X(\bar{\mathbb{K}}) = \varprojlim_\alpha X_\alpha(\bar{\mathbb{K}})$. If (X_α, S) is a uniform algebraic endomotive then

$$\mathcal{P}(X_\alpha, S) = (X, S, \mu)$$

is a measured analytic endomotive with μ the projective limit of the normalized counting measures. On correspondences the functor gives

$$\mathcal{P} : Z \mapsto \mathcal{Z} = Z(\bar{\mathbb{K}}). \tag{4.90}$$

This is compatible with composition of correspondences since in both cases the composition is given by the fiber product over the middle groupoid (cf. (4.54) and (4.68)).

(2) We know by Lemma 4.28 and §2.3 above that the Galois group G acts by automorphisms of the analytic endomotives $\mathcal{P}(X_\alpha, S)$. This action is compatible with the morphisms. This shows that G acts on the functor $\mathcal{P}$ by natural transformations.

(3) The functor $\mathcal{J} : \mathcal{AM}_{\mathbb{K},\mathbb{E}} \to \mathcal{EV}^0_{\mathbb{K},\mathbb{E}}$ determined on $X \in \mathrm{Obj}(\mathcal{V}^0_{\mathbb{K}})$ by $\mathcal{J}(X) = (X_\alpha, S)$, where $X_\alpha = X$ for all α and $S = (1)$ is trivial, gives a faithful embedding of the category of Artin motives in the category of algebraic endomotives. The functor $\mathcal{J}$ is the identity on morphisms.

(4) The functor $k : C^*\mathcal{V}^0_{\mathbb{K},\mathbb{E}} \to \mathcal{KK} \otimes \mathbb{E}$ maps the analytic endomotive $(\mathcal{X}, S)$ to its C^*-algebra $C(\mathcal{X}) \rtimes S$. To see how it acts on morphisms, consider first an étale correspondence $\mathcal{Z}$ as in Definition 4.26. The functor k then assigns to the correspondence $\mathcal{Z}$ the KK-theory class $k(\mathcal{Z}) = [\mathcal{E}_{\mathcal{Z}}, \phi, 0]$ of the bimodule $\mathcal{E}_{\mathcal{Z}}$ (the morphism $\phi : C(\mathcal{X}) \rtimes S \to \mathrm{End}(\mathcal{E}_{\mathcal{Z}})$ is the left $C(\mathcal{X}) \rtimes S$-module structure) with the trivial grading $\gamma = 1$ and the zero endomorphism $F = 0$. Since the correspondence is étale any endomorphism of the right C^*-module $\mathcal{E}_{\mathcal{Z}}$ on $C(\mathcal{X}') \rtimes S'$ is compact. Thus in particular $F^2 - 1 = -1$ is compact. This is exactly the reason why one had to require that the correspondence is étale. This is extended by $\mathbb{E}$-linearity to a map of $M((\mathcal{X}, S), (\mathcal{X}', S')) \otimes_{\mathbb{Q}} \mathbb{E}$ to $\mathcal{KK} \otimes \mathbb{E}$.

Now consider the composite functor

$$k \circ \mathcal{P} \circ \mathcal{J} : \mathcal{AM}_{\mathbb{K},\mathbb{E}} \to \mathcal{KK} \otimes \mathbb{E}.$$

For $X, Y \in \mathrm{Obj}(\mathcal{V}^0_{\mathbb{K}})$, $X = \mathrm{Spec}(A)$ and $Y = \mathrm{Spec}(B)$, two smooth projective zero-dimensional varieties over $\mathbb{K}$, and $U \in \mathrm{Hom}_{\mathcal{AM}_{\mathbb{K},\mathbb{E}}}(X, Y)$ a correspondence, we know by Lemma 4.16 that U is a (virtual) A–B bimodule $\mathcal{E}_U$. Upon changing coefficients to $\mathbb{C}$, one obtains a (virtual) $A_{\mathbb{C}}$–$B_{\mathbb{C}}$ bimodule $\mathcal{E}_U \otimes_{\mathbb{K}} \mathbb{C}$. The algebras are abelian AF algebras of continuous functions on a totally disconnected space. In this case, the points of the space are uniquely specified by the corresponding linear forms $K_0 \to \mathbb{Z}$, and the K_0-group (ordered and pointed) completely determines the algebra. Elements in KK-theory can be specified by homomorphisms of K_0. Thus, such a (virtual) bimodule induces a morphism of their K-theory (tensored with $\mathbb{E}$)

$$U_* : K_0(C(\mathcal{X})) \otimes_{\mathbb{Z}} \mathbb{E} \to K_0(C(\mathcal{X}')) \otimes_{\mathbb{Z}} \mathbb{E}. \tag{4.91}$$

Since the correspondence is determined by the induced morphism

$$K_0(A_{\mathbb{C}}) \otimes \mathbb{E} \to K_0(B_{\mathbb{C}}) \otimes \mathbb{E},$$

we obtain that the composite functor $k \circ \mathcal{P} \circ \mathcal{J}$ is faithful. $\square$

2.6. Self-maps of algebraic varieties.

We now present a large class of examples of uniform algebraic endomotives and corresponding measured analytic endomotives, obtained from self-maps of algebraic varieties.
Let (Y, y_0) denote a pointed algebraic variety over $\mathbb{K}$. Let S be a unital abelian semigroup of self-maps $s : Y \to Y$, with $s(y_0) = y_0$, that are finite, so that they have a well defined degree $\deg(s)$, and unramified over the point y_0. We set

$$X_s = \{y \in Y \mid s(y) = y_0\}. \tag{4.92}$$

Lemma 4.35. *The X_s form a projective system of objects in $\mathcal{V}^0_{\mathbb{K}}$.*

Proof. First notice that the condition that the maps $s \in S$ are finite and unramified at y_0 implies that the preimages X_s are smooth projective zero-dimensional varieties over $\mathbb{K}$.
Suppose given a pair $s, s' \in S$ with $s' = sr$, for some $r \in S$. We then write $s|s'$, meaning that s divides s' in S. For any such pair we have a map

$$\xi_{s,s'} : X_{sr} \to X_s, \quad \text{given by} \quad X_{sr} \ni y \mapsto \xi_{s,s'}(y) = r(y) \in X_s. \tag{4.93}$$

This defines a projective system indexed by the semigroup S itself with partial order given by divisibility. $\square$

Let $X = \varprojlim_{s\in S} X_s$ be the projective limit of a system as above with projection maps $\xi_s : X \to X_s$. We write as before $\mathcal{X} = X(\bar{\mathbb{K}})$ for the set of algebraic points of X. With $X_s = \mathrm{Spec}(A_s)$ and $A = \varinjlim A_s$, an algebraic point of X is a character of A with values in $\bar{\mathbb{K}}$.
Since the self-maps $s \in S$ satisfy $s(y_0) = y_0$, the base point y_0 defines a component Z_s of X_s for all $s \in S$. We denote by $e_s \in A_s$ the corresponding idempotent, and we view it as an idempotent in $A = \varinjlim A_s$. Its image in $A_{s'}$ for $s|s'$ corresponds to the inverse image $\xi^{-1}_{s,s'}(Z_s)$ of Z_s in $X_{s'}$. It is a union of components of $X_{s'}$. This procedure defines, for each $s \in S$, an open and closed subset $\mathcal{X}^{e_s} = X^{e_s}(\bar{\mathbb{K}})$ of the projective limit $\mathcal{X}$, obtained by taking the limit $X^{e_s} = \xi_s^{-1}(Z_s) = \varprojlim_{s'} \xi^{-1}_{s,s'}(Z_s)$. We still denote by e_s the corresponding idempotent, i.e. the characteristic function of X^{e_s}.

Proposition 4.36. *Let $s \in S$. The map β_s defined by*

$$\xi_u \beta_s(x) = s\, \xi_u(x), \quad \forall u \in S\,,\ x \in X, \tag{4.94}$$

is an isomorphism $\beta_s : X \to X^{e_s}$ and the inverse map is given by

$$\xi_u \beta_s^{-1}(x) = \xi_{su}(x), \quad \forall u \in S\,,\ x \in X^{e_s}\,. \tag{4.95}$$

The semigroup S acts on the projective limit X by the partial isomorphisms β_s.

PROOF. Let us check that β_s is well defined as a map $X \to X$ by (4.94). First, with $x_u = \xi_u(x) \in X_u$ one has $sx_u \in X_u$ since $us(x_u) = su(x_u) = s(y_0) = y_0$. Next the $y_u = sx_u \in X_u$ satisfy the compatibility condition $y_u = ry_{ur}$ of the projective system, since $y_u = sx_u = srx_{ur} = rsx_{ur} = ry_{ur}$. The range of β_s is contained in X^{e_s} since by (4.94) we have

$$\xi_s(\beta_s(x)) = s\xi_s(x) = \xi_1(x) = y_0.$$

Let us check similarly that β_s^{-1} is well defined as a map $X^{e_s} \to X$ by (4.95). First, with $x \in X^{e_s}$ one has $\xi_{su}(x) \in X_u$ since $u(x_{su}) = x_s = y_0$. Moreover the $z_u = x_{su}$ satisfy the compatibility condition $z_u = rz_{ur}$ of the projective system, since $z_u = x_{su} = rx_{rsu} = rx_{sru} = rz_{ur}$. Finally when we apply β_s to z we get back x since $sz_u = sx_{su} = x_u$. In the same way when we apply (4.95) to $\beta_s(x)$ we get $z_u = \xi_{su}\beta_s(x) = s\xi_{su}(x) = x_u$ so that (4.95) gives the inverse of β_s. The last statement follows from (4.94). $\square$

At the algebraic level one gets an endomorphism

$$\rho_s \;:\; A \to A_{e_s}\,, \quad \rho_s(f)(x) = f(\beta_s^{-1}(x)) \;\; \forall x \in X^{e_s}\,. \tag{4.96}$$

For $f \in A_u$ one has $\rho_s(f) \in (A_{su})_{e_s}$, which is uniquely determined by the identification

$$X_{su}^{e_s} = \{x \in X_{su} \,:\, ux = y_0\} = \, X_u\,. \tag{4.97}$$

In other words for $f \in A_u$, $\rho_s(f)$ is the element of A_{su} which is zero in the complement of e_s and agrees with f using the identification $X_{su}^{e_s} \sim X_u$ of (4.97).
This result shows that a unital abelian semigroup S of finite endomorphisms of a pointed algebraic variety (Y, y_0), unramified at y_0, determines an algebraic endomotive (X, S). The following result shows that the endomotives constructed in this way are uniform, hence they define corresponding measured analytic endomotives $(\mathcal{X}, S, \mu)$.

LEMMA 4.37. *Suppose given a semigroup of self-maps of* (Y, y_0) *as above and let* X_s *be as in* (4.92). *Let* μ_s *denote the normalized counting measure on the set* $X_s(\bar{\mathbb{K}})$. *These satisfy the compatibility condition*

$$\xi_{s',s}\mu_s = \mu_{s'}, \quad \forall s, s' \in S. \tag{4.98}$$

PROOF. This simply follows from the fact that the self-maps are finite endomorphisms of Y, hence the number of preimages of a point where s is unramified is equal to $\deg s$. $\square$

2.7. The Bost–Connes endomotive.

A main example of an algebraic endomotive and corresponding measured analytic endomotive produced by the method of self-maps of algebraic varieties gives the algebra of the Bost–Connes system [**30**], [**86**] that we already encountered in our discussion of quantum statistical mechanics and $\mathbb{Q}$-lattices in Chapter 3. This example will also play a crucial role in the present chapter, because of its relation to the Riemann zeta function, which we already partly analysed in Chapter 2 and Chapter 3.

THEOREM 4.38. *Let (Y, y_0) be the pointed algebraic variety with Y the affine group scheme $\mathbb{G}_m$ (the multiplicative group) and $y_0 = 1$. Let S be the semigroup of endomorphisms of $\mathbb{G}_m$ of the form $u \mapsto u^n$ for some $n \in \mathbb{N}$. The algebraic endomotive $\mathcal{A}_{\mathbb{Q}} = A \rtimes S$ over $\mathbb{Q}$ obtained by applying the procedure of §2.6 above to these data is the arithmetic subalgebra $\mathcal{A}_{1,\mathbb{Q}}$ of the Bost–Connes system and the corresponding analytic endomotive is the C^*-algebra $\mathcal{A}_1$ of the Bost–Connes system.*

PROOF. Consider the corresponding X_n defined by taking the preimage of the base point 1 as in (4.92). One has by construction $X_n = \mathrm{Spec}(A_n)$ where A_n is the quotient of the algebra $\mathbb{Q}[u(n), u(n)^{-1}]$ of Laurent polynomials by the relation $u(n)^n = 1$. For $n|m$ the map $\xi_{n,m} : X_m \to X_n$ is given by the algebra homomorphism that sends the generator $u(n) \in A_n$ to $u(m)^a \in A_m$ for $a = m/n$. Thus, one obtains an isomorphism

$$\iota : \varinjlim A_n \to \mathbb{Q}[\mathbb{Q}/\mathbb{Z}], \qquad \iota(u(n)) = e(\frac{1}{n}), \tag{4.99}$$

of the inductive limit algebra A with the group ring $\mathbb{Q}[\mathbb{Q}/\mathbb{Z}]$. Here the $\{e(r)\}_{r\in\mathbb{Q}/\mathbb{Z}}$ denote the canonical basis of the group ring.
The algebra $\mathcal{A}_{1,\mathbb{Q}}$ of the Bost–Connes system is the crossed product

$$\mathbb{Q}[\mathbb{Q}/\mathbb{Z}] \rtimes \mathbb{N},$$

where the semigroup action is given on the basis elements $e(r)$ by

$$\rho_n(e(r)) = \frac{1}{n} \sum_{n\ell=r} e(\ell). \tag{4.100}$$

The C^*-algebra $\mathcal{A}_1$ is the crossed product $C^*(\mathbb{Q}/\mathbb{Z}) \rtimes \mathbb{N}$.
To see that we obtain $\mathcal{A}_{1,\mathbb{Q}}$ as the algebraic endomotive $\mathcal{A}_{\mathbb{Q}} = A \rtimes S$ we need to compare the maps ρ_n defined as in (4.94) with the semigroup action (4.100).
We identify the projective limit $\mathcal{X} = X(\bar{\mathbb{Q}})$ with the space of characters

$$\mathcal{X} = \mathrm{Hom}_{\mathrm{Alg}_{\mathbb{Q}}}(\mathbb{Q}[\mathbb{Q}/\mathbb{Z}], \bar{\mathbb{Q}}).$$

The image $\xi_m(x)$ under the projection map $\xi_m : X \to X_m$ is the restriction of $x \in \mathcal{X}$ to the subalgebra $A_m = \mathbb{Q}[\mathbb{Z}/m\mathbb{Z}]$. The endomorphisms ρ_n of (4.100) satisfy $\rho_n(A_k) \subset A_{nk}$ and $\rho_n(1) = e_n \in A_n$, where

$$e_n = \frac{1}{n} \sum_{n\ell=0} e(\ell)\,. \tag{4.101}$$

For $x \in X$ we have $x(e_n) = 1$ iff the restriction $x|_{A_n}$ is the trivial character, i.e. iff $\xi_n(x) = 1$. We have

$$x(\rho_n(e(r))) = \frac{1}{n} \sum_{n\ell=r} x(e(\ell)).$$

This is nonzero iff $x(e_n) = 1$. In that case one finds

$$x(\rho_n(e(r))) = x(e(\ell)), \quad \forall s, \quad \text{with } n\ell = r\,,$$

and in particular

$$x(\rho_n(e(\frac{1}{k}))) = x(e(\frac{1}{nk})). \tag{4.102}$$

When $k|m$ the inclusion $X_k \subset X_m$ of (4.97) is given at the algebra level by the homomorphism

$$j_{k,m} : A_m \to A_k, \quad j_{k,m}(u(m)) = u(k). \tag{4.103}$$

Thus, one can rewrite (4.102) in the form

$$x \circ \rho_n \circ j_{k,nk} = x|_{A_{nk}}. \tag{4.104}$$

Thus, one has

$$\xi_{nk}(x) = \xi_k(x \circ \rho_n),$$

Using (4.95) one then sees that the endomorphisms ρ_n of (4.100) agree with those defined by (4.94), hence the result follows. □

REMARK 4.39. *The nonzero endomorphisms of $\mathbb{G}_m$ correspond to maps of the form $u \mapsto u^n$ for some non-zero $n \in \mathbb{Z}$. To obtain the Bost–Connes system we restrict to the semigroup of positive integers $S = \mathbb{N} \subset \mathbb{Z} \smallsetminus \{0\}$.*

The intertwining properties of extremal KMS_∞ states of the Bost–Connes system discussed in Chapter 3 follows from Proposition 4.30 (up to a normalization replacing $g \mapsto g^{-1}$ in the action of symmetries) and the fact (cf. [**30**]) that these states are induced from pure states of $C(\mathcal{X}) = C^*(\mathbb{Q}/\mathbb{Z})$.

There is an important difference between the description of the Bost–Connes system in terms of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices up to scaling given in Chapter 3 and the one given here as an endomotive. The groups $\mathbb{Q}/\mathbb{Z}$ and $\hat{\mathbb{Z}}$ are Pontrjagin dual, hence one has an isomorphism

$$C^*(\mathbb{Q}/\mathbb{Z}) \simeq C(\hat{\mathbb{Z}}) \tag{4.105}$$

of C^*-algebras realized by the Fourier transform. However, because the Fourier transform is not defined over $\mathbb{Q}$, this is not an isomorphism at the level of the arithmetic structure, i.e. of the underlying algebraic endomotives. In other words, on the one hand we have the system $(X_n, \mathbb{N})$ with $X_n = \mathrm{Spec}(A_n)$ and $A_n = \mathbb{Q}[\mathbb{Z}/n\mathbb{Z}]$ as above and the algebraic endomotive $\mathbb{Q}[\mathbb{Q}/\mathbb{Z}] \rtimes \mathbb{N}$ underlying the analytic endomotive (4.105). On the other hand, we can also construct a system $(X'_n, \mathbb{N})$ with $\mathrm{Spec}(X'_n) = B_n$ with B_n the algebra of $\mathbb{Q}$-valued functions on $\mathbb{Z}/n\mathbb{Z}$. This also gives a $\mathbb{Q}$-algebra $B \rtimes \mathbb{N}$ underlying the same analytic endomotive (4.105). However, these two arithmetic structures are not the same. In fact, in the first case, there is a nontrivial action of $G = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ given on $C(\mathcal{X}) = \mathrm{Hom}(\mathbb{Q}[\mathbb{Q}/\mathbb{Z}], \bar{\mathbb{Q}})$ by composition as in (4.71) and on the system by restricting to the subalgebras A_n of $\mathbb{Q}[\mathbb{Q}/\mathbb{Z}]$. On the other hand, the Galois action is trivial on the B_n.

Here is a brief list of other cases that it would be interesting to investigate in terms of endomotives and self-maps of varieties as in the general procedure of §2.6 above:

- An elliptic curve Y with base point $y_0 = 0$ and with S the abelian semigroup $\mathbb{N}$ of endomorphisms of Y of the form $x \mapsto nx$. The space X then coincides with the total Tate module (cf. Chapter 3 (3.177)).
- An elliptic curve Y with complex multiplication by an imaginary quadratic field $\mathbb{K}$, with base point $y_0 = 0$ and with S the abelian semigroup of nontrivial endomorphisms of Y. The algebraic endomotive $A \rtimes S$, obtained as above, defined over the Hilbert class field of $\mathbb{K}$, gives an arithmetic structure to the C^*-algebra $\mathcal{A}_{1,\mathbb{K}}$ of the system of [**90**], cf. Chapter 3 above. The normal abelian hypothesis of Proposition 4.30 is fulfilled in this case too.

- Endomotives obtained by an analogous procedure from abelian varieties with complex multiplication.
- The quantum statistical mechanical systems for number fields obtained in the context of Shimura varieties in [**159**].
- Some positive characteristic examples, such as the analogs of the Bost–Connes system for function fields of [**172**] and [**104**].
- The space constructed in [**118**] from a characteristic zero lift of an elliptic curve over a finite field with the endomorphism group generated by the Frobenius; see also [**205**].

3. Motives and noncommutative spaces: higher dimensional perspectives

The compatibility of correspondences given by algebraic cycles for motives and by correspondences in the sense of Kasparov modules or of morphisms of cyclic modules for noncommutative spaces is much more subtle in higher dimension than it is for the zero-dimensional case of Artin motives and endomotives. One main reason is that in the zero-dimensional case one does not have to deal with the subtleties involved in the choice of the equivalence relations on cycles. We show here that, in the higher dimensional case there is a natural way to compare correspondences given by algebraic cycles and those given by Kasparov modules defining elements in KK-theory, using a description of KK-theory due to Connes and Skandalis [**98**]. We then discuss various questions related to finding an appropriate setting where motives and noncommutative spaces can be compared.

3.1. Geometric correspondences.

In the topological setting, one considers smooth manifolds X and Y. Recall that the Gysin map $f_! : K^*(X) \to K^*(Y)$ in topological K-theory is defined for K-oriented maps $f : X \to Y$ (cf. [**10**]), where a K-orientation of $f : X \to Y$ is specified by an element of the form $\kappa + 2c$, with $c \in H^2(X, \mathbb{Z})$ and

$$\kappa = w_2(X) - f^* w_2(Y) \mod 2,$$

with w_2 the Stiefel–Whitney class, satisfying $f_!^{\kappa+2c}(E) = f_!^{\kappa}(E \cdot L_c)$, for the line bundle with Chern class $c_1(L_c) = c$.

In fact, as in [**98**], one only needs to assume that Y is a smooth manifold while X is only a locally compact parameter space.

Definition 4.40. *A geometric correspondence is a set of data (Z, E, f_X, g_Y) where Z is a smooth manifold with maps $f_X : Z \to X$ and $g_Y : Z \to Y$, where f_X is continuous and proper and g_Y is continuous and K-oriented, and with E a complex vector bundle over Z.*

Notice that, unlike correspondences given by cycles in the product as in the case of motives, in this setting one does not require that $Z \subset X \times Y$. On the other hand, one has an extra piece of data given by the vector bundle E. (This extra datum should perhaps be compared with some notions of anabelian motives.)

DEFINITION 4.41. *The composition of two given geometric correspondences* (Z_1, E_1, f_X, g_Y) *and* (Z_2, E_2, f_Y, g_W) *is given in this setting by taking the fiber product* $Z = Z_1 \times_Y Z_2$ *and the bundle* $E = \pi_1^* E_1 \times \pi_2^* E_2$*, with* $\pi_i : Z \to Z_i$ *the projections.*

Notice that one needs to assume transversality conditions on the maps g_Y and f_Y in order to ensure that the fibered product is a smooth manifold. In the smooth setting transversality can always be achieved by a small perturbation (cf. §III, [**98**]). This will no longer be the case in the algebraic setting, as we discuss below.

Given X_1 and X_2 smooth manifolds and a continuous oriented map $f : X_1 \to X_2$, there is an element $f! \in KK(X_1, X_2)$ that gives the Grothendieck–Riemann–Roch formula

$$\mathrm{ch}(F \otimes f!) = f_!(\mathrm{Td}(f) \cup \mathrm{ch}(F)), \tag{4.106}$$

for all $F \in K^*(X_1)$, with $\mathrm{Td}(f)$ the Todd genus

$$\mathrm{Td}(f) = \mathrm{Td}(TX_1)/\mathrm{Td}(f^*TX_2). \tag{4.107}$$

Geometric correspondences describe KK-theory classes in the following way (cf. Theorem 3.2 of [**98**]).

PROPOSITION 4.42. [**98**] *The data* (Z, E, f_X, g_Y) *define an element*

$$k(Z, E, f_X, g_Y) = (f_X)_*((E) \otimes (g_Y)!) \tag{4.108}$$

in $KK(X, Y)$*, where* (E) *is the class of* E *in* $KK(Z, Z)$ *and the element* $(g_Y)! \in KK(Z, Y)$ *satisfies* (4.106)*. Moreover, the Kasparov product in* KK*-theory is given by the composition of correspondences*

$$k(Z_1, E_1, f_X, g_Y) \circ k(Z_2, E_2, f_Y, g_W) = k(Z, E, f_X, g_W), \tag{4.109}$$

with $k(Z, E, f_X, g_W) \in KK(X, Z)$.

3.2. Algebraic cycles and K-theory.

In order to compare correspondences given by algebraic cycles with the geometric correspondences described above, as in [**74**], we are going to rephrase the first in K-theoretic rather than (co)homological terms. This can be done by considering the maps induced by algebraic cycles on K-theory (cf. [**214**]).

Let X and Y be smooth projective algebraic varieties and let $Z = \sum_i n_i Z_i$ be an algebraic cycle in $X \times Y$, that is, an element of the free abelian group generated by closed irreducible algebraic subvarieties Z_i of $X \times Y$. (See the section on motives in Chapter 1 above.)

For simplicity, assume that $Z = Z_i$ is an irreducible subvariety of $X \times Y$. We denote by p_X and p_Y the projections of $X \times Y$ onto X and Y, respectively, and we assume that they are proper. We denote by $f_X = p_X|_Z$ and $g_Y = p_Y|_Z$ the restrictions to Z. Notice that, in the case of complex algebraic varieties, the complex structure also determines a K-orientation.

To an irreducible subvariety $T \overset{i}{\hookrightarrow} Y$ we associate the coherent $\mathcal{O}_Y$-module $i_*\mathcal{O}_T$. For simplicity of notation we write it as $\mathcal{O}_T$. We use a similar notation for the coherent sheaf $\mathcal{O}_Z$, for $Z \hookrightarrow X \times Y$ as above.

The map on sheaves that corresponds to the cap product on cocycles is given by

$$Z : \mathcal{O}_T \mapsto (p_X)_* \left(p_Y^* \mathcal{O}_T \otimes_{\mathcal{O}_{X \times Y}} \mathcal{O}_Z \right). \tag{4.110}$$

The result is a coherent sheaf, since p_X is proper.

The functor $f_!$ is right adjoint to f^*, i.e. $f^* f_! = \mathrm{id}$, and that it satisfies the Grothendieck–Riemann–Roch formula

$$\mathrm{ch}(f_!(F)) = f_!(\mathrm{Td}(f) \cup \mathrm{ch}(F)). \tag{4.111}$$

We have the following compatibility result between algebraic cycles and geometric correspondences.

LEMMA 4.43. *To a correspondence of the form* (4.110) *one can associate a geometric correspondence* $\mathcal{P}(Z) = (Z, E, f_X, g_Y)$ *in such a way that the composition (for cycles in general position) is compatible with the composition of geometric correspondences.*

PROOF. First notice that, using (4.111), we can equally compute the intersection product of (4.110) by first computing

$$\mathcal{O}_T \otimes_{\mathcal{O}_X} (p_Y)_! \mathcal{O}_Z \tag{4.112}$$

and then applying p_Y^*. Using (4.111) and (4.106) we know that we can replace (4.112) by $\mathcal{O}_T \otimes_{\mathcal{O}_X} (\mathcal{O}_Z \otimes (p_Y)!)$ with the same effect on K-theory.

Thus, to a correspondence in the sense of (4.110) given by an algebraic cycle $Z \subset X \times Y$ we associate the geometric correspondence (Z, E, f_X, g_Y) with $f_X = p_X|_Z$ and $g_Y = p_Y|_Z$ and with the bundle $E = \mathcal{O}_Z$.

Now we consider the composition of correspondences. Suppose given smooth projective varieties X, Y, and W, and (virtual) correspondences $U = \sum a_i Z_i$ and $V = \sum c_j Z'_j$, with $Z_i \subset X \times Y$ and $Z'_j \subset Y \times W$ closed reduced irreducible subschemes. The composition of correspondences is then given in terms of the intersection product

$$U \circ V = (\pi_{13})_*((\pi_{12})^* U \bullet (\pi_{23})^* V), \tag{4.113}$$

with the projection maps $\pi_{12} : X \times Y \times W \to X \times Y$, $\pi_{23} : X \times Y \times W \to Y \times W$, and $\pi_{13} : X \times Y \times W \to X \times W$. For simplicity, again assume that U is given by a single subvariety $Z_1 \subset X \times Y$ and V by a single $Z_2 \subset Y \times W$.

We are assuming a transversality condition in the algebraic setting, namely the hypothesis that $(\pi_{12})^* Z_1$ and $(\pi_{23})^* Z_2$ are in general position in $X \times Y \times W$.

Recall that, if two closed subschemes $T_1, T_2 \subset X$ are in general position, and $\mathcal{O}_{T_i}$ denote the corresponding structure sheaves, then their product in $K_*(X)$ agrees with the intersection product ([**214**], Theorem 2.3)

$$[\mathcal{O}_{T_1}] \cdot [\mathcal{O}_{T_2}] = [\mathcal{O}_{T_1 \bullet T_2}]. \tag{4.114}$$

Using (4.114) we write the composition (4.113) in the form

$$Z_1 \circ Z_2 = (\pi_{13})_*([\pi_{12}^* \mathcal{O}_{Z_1}] \cdot [\pi_{23}^* \mathcal{O}_{Z_2}]). \tag{4.115}$$

Notice that, for $Z_1 \subset X \times Y$ and $Z_2 \subset Y \times W$ the intersection in $X \times Y \times W$ is exactly the fiber product considered in [**98**], $\pi_{12}^{-1} Z_1 \cap \pi_{23}^{-1} Z_2 = Z = Z_1 \times_Y Z_2$, with $\mathcal{O}_Z = \pi_{12}^* \mathcal{O}_{Z_1} \otimes \pi_{23}^* \mathcal{O}_{Z_2}$. Moreover, we have $\pi_{13} = (f_X, g_W) : Z \to X \times W$, where both f_X and g_W are proper, so that we can identify $(f_X)_*(\mathcal{O}_Z \otimes (g_W)!)$ with $(\pi_{13})_*(\pi_{12}^* \mathcal{O}_{Z_1} \otimes \pi_{23}^* \mathcal{O}_{Z_2})$ as desired. In fact, Theorem 3.2 of [**98**] shows that $(f_X)_*(\mathcal{O}_Z \otimes (g_W)!)$ represents the Kasparov product. $\square$

REMARK 4.44. *In the case of the product of cycles that are not in general position the formula* (4.114) *is modified by Tor corrections and one obtains (*[**214**]*, Theorem 2.7)*

$$[\mathcal{O}_{T_1}]\cdot[\mathcal{O}_{T_2}] = \sum_{i=0}^{n}(-1)^i\left[\mathrm{Tor}_i^{\mathcal{O}_X}(\mathcal{O}_{T_1},\mathcal{O}_{T_2})\right]. \tag{4.116}$$

3.2.1. *Algebraic versus topological K-theory.*

We have encountered so far different type of correspondences: Kasparov modules and their KK-theory classes, algebraic cycles and their K-theoretic version, morphisms in the category of cyclic modules. It is clearly necessary to establish more precise relations between these different settings.

A first observation is that often in cases of interest (like the scaling action that we will concentrate on later in this chapter) the "primary" invariants obtained by the Chern character map from K-theory to cyclic cohomology are trivial (in the case of the scaling action because it is an action of a continuous group). Thus, to accommodate such cases, one needs to develop a setting where secondary invariants also appear.

Moreover, the comparison described above between the correspondences given by algebraic cycles (viewed through the K-theoretic description) and the geometric correspondences and their induced KK-theory classes shows that one needs a more refined setting that combines algebraic and topological K-theory.

It seems to be relevant to this effect (for both of the considerations listed above) to recall a result of Connes and Karoubi [**81**] that gives a relative version in the fibration relating algebraic and topological K-theory.

PROPOSITION 4.45. [**81**] *By viewing both algebraic and topological K-theory in terms of homotopy groups of corresponding classifying spaces, one obtains a fibration and a corresponding long exact sequence. This sequence is related to the long exact sequence for Hochschild and cyclic cohomology of* [**67**] *through the Chern character ch_* and a regulator map D_*, so that one has a commutative diagram with exact rows*

$$\begin{tikzcd}
K^{\mathrm{top}}_{n+1}(A) \arrow[r] \arrow[d, "ch_{n+1}"] & K^{\mathrm{rel}}_n(A) \arrow[r] \arrow[d, "ch^{\mathrm{rel}}_n"] & K^{\mathrm{alg}}_n(A) \arrow[r] \arrow[d, "D_n/n!"] & K^{\mathrm{top}}_n(A) \arrow[r] \arrow[d, "ch_n"] & K^{\mathrm{rel}}_{n-1}(A) \arrow[d, "ch^{\mathrm{rel}}_{n-1}"] \\
HC_{n+1}(A) \arrow[r, "S"] & HC_{n-1}(A) \arrow[r, "B/n"] & HH_n(A) \arrow[r, "I"] & HC_n(A) \arrow[r, "S"] & HC_{n-2}(A).
\end{tikzcd}$$

In a related perspective, motivic cohomology can be described in terms of the higher Chow groups of Bloch (cf. [**22**], [**286**]). It was shown in [**238**] that one can formulate a version of the higher Chow group complex that is modelled on the (b, B)-bicomplex of cyclic cohomology and obtain an analog of the exact sequence relating Hochschild and cyclic cohomology.

3.2.2. *Abelian varieties and noncommutative tori.*

A good setting in which to see an interesting interplay between motives and noncommutative spaces that goes beyond the zero-dimensional case of endomotives will be obtained by looking at the case of curves.

Motives are designed to be a universal cohomology theory. Since all cohomologies for a smooth projective curve factor through the Jacobian of the curve, it is clear that pure motives of curves are Jacobians. However, when one considers the subcategory of the category of pure motives spanned by the motives of curves, one also finds more general abelian varieties. In fact, the subcategory of the category of pure motives generated by the motives of curves is that of abelian varieties.

In the simplest example of the elliptic curves, it is possible to view the noncommutative tori of [**64**] as degenerations of elliptic curves, by viewing an elliptic curve through its analytic model (Jacobi uniformization) $E = \mathbb{C}^*/q^{\mathbb{Z}}$ for $|q| < 1$ and taking a limit $q \to e^{2\pi i\theta}$ with $\theta \in \mathbb{R} \smallsetminus \mathbb{Q}$ (cf. §2.3 of [**221**] and [**219**]). At the algebraic level, noncommutative tori can be described through the algebraic noncommutative space obtained by Polishchuk [**244**] from the holomorphic bundles of [**96**].

Thus, it seems interesting to extend the category of motives of abelian varieties by at least including noncommutative spaces such as (higher-rank) noncommutative tori. Notice in particular that there is no good construction of noncommutative spaces (in the analytic sense) that correspond to higher genus Riemann surfaces. A motivic viewpoint can in fact solve this problem by defining the motives of noncommutative curves (i.e. their Jacobians). This amounts to an analog for noncommutative tori of the classical Schottky problem of recognizing Jacobians among abelian varieties. The Schottky problem admits a solution that completely characterizes Jacobians in terms of a hierarchy of differential equations (KP soliton equations) satisfied by the theta functions, [**269**]. This hints at a possible characterization of motives of noncommutative curves in terms of theta functions of noncommutative tori (cf. e.g. [**215**]).

4. A thermodynamic "Frobenius" in characteristic zero

We have seen in Section 2 above that the data of a uniform algebraic endomotive (X_α, S) over $\mathbb{K}$ yield the following objects:

- A C^*-algebra $\mathcal{A} = C(\mathcal{X}) \rtimes S$ of the corresponding analytic endomotive.
- An action of the Galois group $G = \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})$ on it.
- An arithmetic subalgebra $\mathcal{A}_{\mathbb{K}} = A \rtimes S$ given by the algebraic endomotive.
- A state φ on $\mathcal{A}$, induced by the measure μ on $\mathcal{X}$.

In this section we give a general construction, obtained using operator algebra techniques and quantum statistical mechanics, that starting from the data of a unital involutive algebra $\mathcal{A}_{\mathbb{C}}$ and a state φ on a C^*-completion $\mathcal{A}$ of $\mathcal{A}_{\mathbb{C}}$ yields a cohomology with a representation of the multiplicative group $\mathbb{R}_+^*$. In this general procedure the algebra need not be an endomotive (in particular the procedure applies as well to higher dimensional cases), but we will be especially interested in applying it to endomotives and in particular to the case of the Bost–Connes algebra.
Here are some main features of the construction we describe in this section applied to endomotives in comparison to the case of curves over $\mathbb{F}_q$:

- Passing to the "dual system" of a quantum statistical mechanical system associated to the data $(\mathcal{A}, \varphi)$ gives an analog in characteristic zero of passing from $\mathbb{F}_q$ to its algebraic closure $\bar{\mathbb{F}}_q$.

- Taking a cokernel $D(\mathcal{A},\varphi)$ of a suitable map of noncommutative spaces in the abelian category of Λ-modules (a procedure we call “distillation”) yield something comparable to a motivic $H^1_{\rm abs}$ (cf. (4.28) and (4.29) above).
- The scaling action of $\mathbb{R}_+^*$ on cyclic homology $HC_0(D(\mathcal{A},\varphi))$ provides a replacement for the action of Frobenius on ℓ-adic cohomology.

We discuss these various aspects in detail in the remainder of this chapter.

4.1. Tomita's theory and the modular automorphism group.

Let $\mathcal{A}_{\mathbb{C}}$ be a unital involutive algebra over $\mathbb{C}$. We assume that $\mathcal{A}_{\mathbb{C}}$ has a countable basis (as a vector space over $\mathbb{C}$) and that the expression

$$\|a\| = \sup \|\pi(a)\|, \tag{4.117}$$

with π ranging over all unitary Hilbert space representations of $\mathcal{A}_{\mathbb{C}}$, defines a *finite* norm on $\mathcal{A}_{\mathbb{C}}$. We denote by $\mathcal{A}$ the norm completion of $\mathcal{A}_{\mathbb{C}}$ with respect to (4.117). It is a unital C^*-algebra.

For instance, this applies to the case where $\mathcal{A}_{\mathbb{C}} = \mathcal{A}_{\mathbb{K}} \otimes_{\mathbb{K}} \mathbb{C}$ is obtained from an algebraic endomotive defined over a number field $\mathbb{K}$, with the resulting C^*-algebra given by the analytic endomotive $C(\mathcal{X}) \rtimes S$.

Given a state $\varphi : \mathcal{A} \to \mathbb{C}$, i.e. a linear form such that $\varphi(1) = 1$ and $\varphi(a^*a) \geq 0$, for all $a \in \mathcal{A}$, we consider, as in Proposition 3.5 of Chapter 3, the associated GNS representation π_φ with Hilbert space $\mathcal{H}_\varphi$, given by the completion of $\mathcal{A}/\mathcal{I}_\varphi$ in the sesquilinear form

$$\langle a, b\rangle = \varphi(a^*b), \quad \forall a, b \in \mathcal{A}, \tag{4.118}$$

with $\mathcal{I}_\varphi = \{a : \varphi(a^*a) = 0\}$. The vector $\xi = 1 + \mathcal{I}_\varphi$ is cyclic separating for $\pi_\varphi(\mathcal{A})$. We let $\mathcal{M}_\varphi$ denote the von Neumann algebra obtained as the weak closure of $\pi_\varphi(\mathcal{A})$, i.e. the double commutant

$$\mathcal{M}_\varphi = \pi_\varphi(\mathcal{A})''. \tag{4.119}$$

In particular, we consider states with the following property.

DEFINITION 4.46. *A state φ on a C^*-algebra $\mathcal{A}$ is* regular *if the vector $\xi = 1+\mathcal{I}_\varphi \in \mathcal{H}_\varphi$ is cyclic also for the commutant $\pi_\varphi(\mathcal{A})'$.*

For regular states, the subspaces $\mathcal{M}_\varphi\xi \subset \mathcal{H}_\varphi$ and $\mathcal{M}'_\varphi\xi \subset \mathcal{H}_\varphi$ are dense in $\mathcal{H}_\varphi$ and one can define antilinear operators

$$\begin{aligned} S_\varphi &: \mathcal{M}_\varphi\xi \to \mathcal{M}_\varphi\xi, \quad a\xi \mapsto S_\varphi(a\xi) = a^*\xi \\ S_\varphi^* &: \mathcal{M}'_\varphi\xi \to \mathcal{M}'_\varphi\xi, \quad a'\xi \mapsto S_\varphi^*(a'\xi) = a'^*\xi. \end{aligned} \tag{4.120}$$

These are closable operators, hence they admit a polar decomposition

$$S_\varphi = J_\varphi \Delta_\varphi^{1/2}, \tag{4.121}$$

where J_φ is a conjugate linear involution satisfying

$$J_\varphi = J_\varphi^* = J_\varphi^{-1} \tag{4.122}$$

and $\Delta_\varphi = S_\varphi^* S_\varphi$ is a positive self-adjoint operator satisfying

$$J_\varphi \Delta_\varphi J_\varphi = S_\varphi S_\varphi^* = \Delta_\varphi^{-1}. \tag{4.123}$$

An important result of Tomita [**279**] (cf. Appendix §9.3) shows that the following two properties hold:

$$\begin{aligned} J_\varphi \mathcal{M}_\varphi J_\varphi &= \mathcal{M}'_\varphi, \\ \Delta^{it}_\varphi \mathcal{M}_\varphi \Delta^{-it}_\varphi &= \mathcal{M}_\varphi. \end{aligned} \tag{4.124}$$

In particular this makes it possible to introduce the *modular automorphism group*

$$\sigma^\varphi : \mathbb{R} \to \mathrm{Aut}(\mathcal{M}_\varphi), \qquad \sigma^\varphi_t(a) = \Delta^{-it}_\varphi a \Delta^{it}_\varphi. \tag{4.125}$$

The next important result, which is due to Takesaki, is the fact that the state φ satisfies the KMS_1 condition with respect to the time evolution σ^φ_t. We reproduce here the simple proof of this fact.

LEMMA 4.47. *The state φ satisfies the KMS_1 condition for the time evolution σ^φ_t.*

PROOF. We can write the state in the form $\varphi(a) = \langle \xi, a\xi \rangle$ with $\xi = 1 + \mathcal{I}_\varphi$ in $\mathcal{H}_\varphi$. Moreover, we have $\Delta^z_\varphi \xi = \xi$. We then get, assuming for simplicity that $z \mapsto \sigma^\varphi_z(b)$ extends to the strip $\Im z \in [0,1]$,

$$\begin{aligned} \varphi(\sigma^\varphi_t(b)a) &= \langle \xi, b\Delta^{it}_\varphi a\xi \rangle = \langle b^*\xi, \Delta^{it}_\varphi a\xi \rangle = \langle J\Delta^{1/2}_\varphi b\xi, J\Delta^{it}_\varphi \Delta^{1/2}_\varphi a^*\xi \rangle \\ &= \langle \Delta^{it}_\varphi \Delta^{1/2}_\varphi a^*\xi, \Delta^{1/2}_\varphi b\xi \rangle = \langle \xi, a\Delta^{1-it}_\varphi b\xi \rangle = \varphi(a\sigma^\varphi_{t+i}(b)). \end{aligned}$$

The general case follows by weak density of the algebra of elements b fulfilling the above analyticity requirement for the one-parameter group σ_t. It gives the KMS_1 condition of Definition 3.6 of Chapter 3. ☐

Thus, we have seen that, given the data of a C^*-algebra $\mathcal{A}$ and a state φ, one obtains a natural time evolution σ^φ_t on the von Neumann algebra $\mathcal{M}_\varphi$. In general this need not preserve the C^*-algebra $\mathcal{A}$ nor the dense subalgebra $\mathcal{A}_\mathbb{C}$ we began with, so we make the following additional requirement on the data $(\mathcal{A}, \varphi)$.

REMARK 4.48. *We restrict to consider cases where the time evolution σ^φ_t of the modular automorphism group on $\mathcal{M}_\varphi$ globally preserves the unital involutive algebra $\mathcal{A}_\mathbb{C}$ and the C^*-algebra $\mathcal{A}$.*

Let us follow all this in the specific example of the Bost–Connes endomotive of §2.7.

LEMMA 4.49. *Let $\mathcal{A}_\mathbb{Q} = A \rtimes S$ be the unital involutive algebra of the Bost-Connes (BC) algebraic endomotive, as in Theorem 4.38, and let $\mathcal{A} = C(\hat{\mathbb{Z}}) \rtimes \mathbb{N}$ be the BC C^*-algebra of the corresponding analytic endomotive. The state φ on $\mathcal{A}$ induced by the measure μ on $\hat{\mathbb{Z}}$ is of the form*

$$\varphi(f) = \int_{\hat{\mathbb{Z}}} f(1,\rho)\, d\mu(\rho), \tag{4.126}$$

and the modular automorphism group restricts to the C^*-algebra as the time evolution of the BC system.

PROOF. We identify elements in the algebra $\mathcal{A}$ with functions $f(g,\rho)$ of elements $\rho \in \hat{\mathbb{Z}}$ and $g \in \mathbb{Q}^*_+$ such that $g\rho \in \hat{\mathbb{Z}}$. We can extend the functions by zero to all of $\mathbb{Q}^*_+ \times \hat{\mathbb{Z}}$ so as to write the product in $\mathcal{A}$ simply in the form

$$f_1 * f_2(g,\rho) = \sum_r f_1(gr^{-1}, g\rho) f_2(r,\rho).$$

The time evolution of the BC system (cf. Chapter 3) is of the form

$$\sigma_t(f(r,\rho)) = r^{it} f(r,\rho). \tag{4.127}$$

Consider now the state on $\mathcal{A}$ defined by (4.126). The measure μ is the projective limit measure on $\hat{\mathbb{Z}} = \varprojlim_n \mathbb{Z}/n\mathbb{Z}$ induced by the normalized uniform counting measures on $\mathbb{Z}/n\mathbb{Z}$. This satisfies the scaling property

$$d\mu(r\rho) = r^{-1} d\mu(\rho), \tag{4.128}$$

for $\rho \in \hat{\mathbb{Z}}$ and $r \in \mathbb{Q}_+^*$ with $r\rho \in \hat{\mathbb{Z}}$. The state (4.126) is KMS_1 with respect to the time evolution (4.127). In fact, we have

$$\varphi(f_1 * f_2) = \sum_r \int_{\hat{\mathbb{Z}}} f_1(r^{-1}, r\rho) f_2(r,\rho)\, d\mu(\rho)$$

$$= \sum_r \int_{\hat{\mathbb{Z}}} f_2(r^{-1}, r\rho) f_1(r,\rho)\, r^{-1}\, d\mu(\rho) = \varphi(f_2 * \sigma_i(f_1)).$$

It follows that σ_t extends to the one-parameter group of modular automorphisms σ_t^φ of the von Neumann algebra obtained from the GNS representation $\mathcal{H}_\varphi$ of the state (4.126). We also refer to [**30**] where it was proved that (4.127) is indeed the modular automorphism group associated to the regular representation of the Hecke algebra $\mathcal{A}$. The space $\mathcal{H}_\varphi$ of the state (4.126) is in fact the Hilbert space of the regular representation. In particular, in this case the modular automorphism group preserves the C^*-algebra $\mathcal{A}$ and the dense subalgebra $\mathcal{A}_{\mathbb{C}}$. The state (4.127) is of type III_1 (cf. [**30**]). ☐

Under the assumption of Remark 4.48, we have a time evolution σ_t^φ on the C^*-algebra $\mathcal{A}$ and it is then a natural question to look at the structure of KMS_β states for this same time evolution, but at other values of the thermodynamic parameter β.

4.2. Regular extremal KMS states (cooling).

We have seen in the various cases of quantum statistical mechanical systems analyzed in Chapter 3 that typically one finds that, for sufficiently large values of β (sufficiently small temperatures) the system resembles more closely a classical space, in the sense that the KMS states pass from being type III at high temperatures (a very "non-classical" case) to being type I_∞ at sufficiently low temperature. While from the point of view of von Neumann algebra the type III case is more interesting, we have seen that, from a more number theoretic viewpoint, the low temperature states tend to have interesting values on suitable arithmetic subalgebras and a more interesting action of the group of symmetries of the system. Thus, it makes sense in our present context to investigate the KMS_β states for $(\mathcal{A}, \sigma_t^\varphi)$ for values $\beta > 1$. This general picture brings us to further restrict the class of data $(\mathcal{A}, \varphi)$ by an additional requirement on the existence of a phase transition from the typically type III state φ at $\beta = 1$ to type I_∞ states at sufficiently large $\beta > 1$.

Remark 4.50. *In the following we restrict our attention to data $(\mathcal{A}, \varphi)$ with the property that, for sufficiently large β, all extremal KMS_β states for the time evolution σ_t^φ on $\mathcal{A}$ are type I_∞ factor states.*

This is the case for instance in the example of the BC system, by the analysis of KMS states of [**30**] discussed in Chapter 3. Physically, the assumption of Remark 4.50 can be regarded as a "typical behavior" of a quantum statistical mechanical system, exhibiting a more chaotic equilibrium state at high temperatures and more ordered phases at lower temperatures.

DEFINITION 4.51. *Let $\mathcal{E}_\beta$ denote the set of extremal KMS$_\beta$ states for the system $(\mathcal{A},\sigma_t^\varphi)$. We denote by $\Omega_\beta \subset \mathcal{E}_\beta$ the set of those extremal states that are of type I_∞ and are regular in the sense of Definition 4.46 above.*

The following result shows that states in Ω_β can all be written in the Gibbs form (see (4.129)). This implies in particular that such states continue to exist at any lower temperature, as will be shown in Lemma 4.56 below.

PROPOSITION 4.52. *Let $\varepsilon \in \Omega_\beta$ be a regular extremal KMS$_\beta$ state of type I_∞ on $(\mathcal{A},\sigma_t^\varphi)$. Then there exists an irreducible representation π_ε of $\mathcal{A}$ on a Hilbert space $\mathcal{H}_\varepsilon^{(1)}$ and an unbounded self-adjoint operator H on $\mathcal{H}_\varepsilon^{(1)}$, with $e^{-\beta H}$ of trace class, such that*

$$\varepsilon(a) = \frac{\mathrm{Tr}(\pi_\varepsilon(a)\, e^{-\beta H})}{\mathrm{Tr}(e^{-\beta H})}, \qquad \forall a \in \mathcal{A}. \tag{4.129}$$

The operator H is the infinitesimal generator of the time evolution in the representation, namely

$$e^{itH}\, \pi_\varepsilon(a)\, e^{-itH} = \pi_\varepsilon(\sigma_t^\varphi(a)), \qquad \forall a \in \mathcal{A}. \tag{4.130}$$

PROOF. Let $\xi_\varepsilon = 1 + \mathcal{I}_\varepsilon \in \mathcal{H}_\varepsilon$ be the cyclic and separating vector in the GNS representation associated to the regular state $\varepsilon \in \Omega_\beta$. The von Neumann algebra $\mathcal{M}_\varepsilon$ generated by $\mathcal{A}$ is a type I_∞ factor. This implies that one has a factorization

$$\mathcal{H}_\varepsilon = \mathcal{H}_\varepsilon^{(1)} \otimes \mathcal{H}_\varepsilon^{(2)}, \quad \text{such that} \quad \mathcal{M}_\varepsilon = \{T \otimes 1 | T \in \mathcal{L}(\mathcal{H}_\varepsilon^{(1)})\}. \tag{4.131}$$

We let π_ε denote the representation $\pi_\varepsilon : \mathcal{A} \to \mathcal{L}(\mathcal{H}_\varepsilon^{(1)})$ induced by the action of $\mathcal{M}_\varepsilon$ in the factorization (4.131), by $a \mapsto \pi_\varepsilon(a) \otimes 1$. The representation π_ε is irreducible since $\pi_\varepsilon(\mathcal{A})$ is weakly dense in $\mathcal{L}(\mathcal{H}_\varepsilon^{(1)})$. The state ε on $\mathcal{A}$ extends to a state

$$\varepsilon(a) = \langle \xi_\varepsilon, a\, \xi_\varepsilon \rangle_{\mathcal{H}_\varepsilon}, \quad \forall a \in \mathcal{M}_\varepsilon \tag{4.132}$$

which is faithful since ξ_ε is cyclic for the commutant $\mathcal{M}_\varepsilon'$ and is normal by construction, hence it can be uniquely written in the form

$$\varepsilon(T \otimes 1) = \langle (T \otimes 1)\xi_\varepsilon, \xi_\varepsilon \rangle_{\mathcal{H}_\varepsilon} = \mathrm{Tr}_{\mathcal{H}_\varepsilon^{(1)}}(T\rho), \tag{4.133}$$

for a density matrix ρ, that is, a positive trace class operator on $\mathcal{H}_\varepsilon^{(1)}$ with $\mathrm{Tr}(\rho) = 1$. We can write the density matrix in the form

$$\rho = \frac{e^{-\beta H}}{\mathrm{Tr}(e^{-\beta H})}$$

for some self-adjoint operator H with $\mathrm{Tr}(e^{-\beta H}) < \infty$. Thus, (4.133) gives (4.129) and it also shows that the operator H has discrete spectrum bounded below. The normal state ε on the type I factor $\mathcal{M}_\varepsilon$ then clearly satisfies the KMS$_\beta$ condition with respect to the time evolution defined by

$$\sigma_t^{(1)}(a) := e^{itH}\, a\, e^{-itH}, \quad \forall a \in \mathcal{M}_\varepsilon.$$

We need to show that this agrees with the extension by continuity of the time evolution σ_t^φ to $\mathcal{M}_\varepsilon$. Let $\tilde{H}$ be an infinitesimal generator for the time evolution σ_t^φ in the GNS representation on $\mathcal{H}_\varepsilon$, which makes sense since ε is invariant under σ_t^φ. The factor $\mathcal{M}_\varepsilon$ is the weak closure of the C^*-algebra $\mathcal{A}$ in $\mathcal{L}(\mathcal{H}_\varepsilon)$ of the GNS representation, hence the extension (4.132) of the state ε to $\mathcal{M}_\varepsilon$ still satisfies the KMS_β condition for the time evolution

$$\tilde{\sigma}_t(T \otimes 1) = e^{it\tilde{H}}\,(T \otimes 1)\,e^{-it\tilde{H}}, \quad \forall (T \otimes 1) \in \mathcal{M}_\varepsilon$$

that extends σ_t^φ to $\mathcal{M}_\varepsilon$. This implies, by [**279**] Theorem 13.2, that the two time evolutions $\sigma_t^{(1)}$ and $\tilde{\sigma}_t = \sigma_t^\varphi$ agree on $\pi_\varepsilon(\mathcal{A})$. □

REMARK 4.53. *As above let π_ε be the irreducible representation of $\mathcal{A}$ associated to a regular extremal state of type I_∞, $\varepsilon \in \Omega_\beta$. The Hamiltonian H implementing the time evolution σ_t^φ as in* (4.130) *is not unique. It is only determined up to an additive constant*

$$H \mapsto H + \log\lambda, \qquad \text{for} \quad \lambda \in \mathbb{R}_+^*. \tag{4.134}$$

REMARK 4.54. *The states $\varepsilon \in \Omega_\beta$ are uniquely determined by the corresponding representation π_ε, due to the formula* (4.129), *which does not depend on the additional choice of H fulfilling* (4.130).

Instead of the space Ω_β of regular extremal KMS_β states of type I_∞, it is in fact better to work with a principal $\mathbb{R}_+^*$-bundle (equivalently a real line bundle if written additively) over Ω_β, which takes into account this ambiguity in the choice of a generator of the time evolution in the corresponding irreducible representation.

DEFINITION 4.55. *The space $\tilde{\Omega}_\beta$ is given by all pairs*

$$\tilde{\Omega}_\beta = \{(\varepsilon, H) \,|\, \varepsilon \in \Omega_\beta, \ \ \pi_\varepsilon(\sigma_t^\varphi(\cdot)) = e^{itH}\pi_\varepsilon(\cdot)e^{-itH}\}. \tag{4.135}$$

It is a principal $\mathbb{R}_+^*$-bundle $\mathbb{R}_+^* \to \tilde{\Omega}_\beta \xrightarrow{p} \Omega_\beta$ with the action

$$\lambda\,(\varepsilon, H) = (\varepsilon, H + \log\lambda) \quad \forall \lambda \in \mathbb{R}_+^* \tag{4.136}$$

The choice of a section by the condition $\mathrm{Tr}(e^{-\beta H}) = 1$ gives a trivialization

$$\tilde{\Omega}_\beta \simeq \Omega_\beta \times \mathbb{R}_+^*.$$

Finally, notice that the properties of type I_∞ regular extremal KMS states are "stable" with respect to further lowering the temperature. In general, the size of Ω_β is a non-decreasing function of β, as the following result shows.

LEMMA 4.56. *For all $\beta' > \beta$, the formula* (4.129) *defines a canonical injection*

$$c_{\beta',\beta} : \Omega_\beta \to \Omega_{\beta'} \tag{4.137}$$

which extends to an $\mathbb{R}_+^*$-equivariant map of the bundles $\tilde{\Omega}_\beta \to \tilde{\Omega}_{\beta'}$.

PROOF. The condition $\mathrm{Tr}(e^{-\beta H}) < \infty$ implies that, for all $\beta' > \beta$, we also have $\mathrm{Tr}(e^{-\beta' H}) < \infty$. Thus, the map $c_{\beta',\beta}$ is well defined. It is straightforward that the state ψ given by (4.129) for β' is $\mathrm{KMS}_{\beta'}$. Notice that it is the restriction of the faithful normal state on $\mathcal{M}_\varepsilon = \mathcal{L}(\mathcal{H}_\varepsilon^{(1)})$ which is given by the density matrix

$$\rho' = \frac{e^{-\beta' H}}{\mathrm{Tr}(e^{-\beta' H})} = \lambda\, \rho^{\beta'/\beta}$$

with ρ as in (4.133) and λ a normalization factor. One can thus obtain the GNS construction for the state ψ on $\mathcal{A}$ as follows. The representation is $a \mapsto \pi_\varepsilon(a) \otimes 1$ in the Hilbert space (4.131) which can be identified with the Hilbert space of Hilbert-Schmidt operators in $\mathcal{H}_\varepsilon^{(1)}$. The cyclic separating vector ξ_ψ of the GNS construction is given by the Hilbert–Schmidt operator which is the positive square root of ρ'. Since the weak closure of $\mathcal{A}$ is the type I factor $\mathcal{M}_\varepsilon$, it follows that ψ is extremal of type I_∞ and is regular since the extended state is faithful on $\mathcal{M}_\varepsilon$. The map $c_{\beta',\beta}$ is injective, since one recovers the original KMS_β state by the reverse procedure. The latter in general only makes sense for those values of β for which $\mathrm{Tr}(e^{-\beta H}) < \infty$. This set of values is a half line with lower bound depending on the corresponding element of $\Omega_\infty := \varinjlim_\beta \Omega_\beta$. $\square$

In [**74**] we refer to the process that assigns to the data $(\mathcal{A}, \varphi)$ the space Ω_β (or Ω_∞) as the "cooling process" (cf. also §4.8 below). It can be thought of as a way of extracting the "classical points" of the noncommutative system $(\mathcal{A}, \sigma_t^\varphi)$ by lowering the temperature and considering the resulting thermodynamical equilibrium states.

4.3. The dual system.

We now present another key step which is a general construction that assigns to a quantum statistical mechanical system $(\mathcal{A}, \sigma_t)$ a dual system $(\hat{\mathcal{A}}, \theta_\lambda)$, where the algebra of the dual system is obtained by taking the crossed product of $\mathcal{A}$ by the time evolution and it comes endowed with a scaling action $\theta_\lambda : \mathbb{R}_+^* \to \mathrm{Aut}(\hat{\mathcal{A}})$. This general procedure originates from the theory of von Neumann algebras and factors as a duality between type III factors and type II factors with automorphisms scaling the trace introduced in Connes' thesis [**63**] in the classification of type III factors and formalized in general terms by Takesaki in [**280**] (cf. also [**99**]).

Since we are especially interested in data $(\mathcal{A}, \sigma_t)$ that come from an underlying uniform algebraic endomotive, we consider the case where the C^*-algebra $\mathcal{A}$ is the norm completion of a dense subalgebra $\mathcal{A}_{\mathbb{C}}$ with countable basis, which is preserved by the time evolution, and we construct the dual system at the level of the dense subalgebra.

DEFINITION 4.57. *The Schwartz space $\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ is the direct limit*

$$\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}}) = \varinjlim_V \mathcal{S}(\mathbb{R}, V). \tag{4.138}$$

of the Schwartz spaces $\mathcal{S}(\mathbb{R}, V)$ of finite dimensional *subspaces $V \subset \mathcal{A}_{\mathbb{C}}$.*

We assume as before that the unital algebra $\mathcal{A}_{\mathbb{C}}$ has a countable basis as a $\mathbb{C}$-vector space. The Schwartz space $\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ has the following algebra structure induced by the action of the time evolution on $\mathcal{A}_{\mathbb{C}}$.

LEMMA 4.58. *The Schwartz space $\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ is an algebra under the product*

$$(x \star y)(s) = \int_{\mathbb{R}} x(t)\sigma_t(y(s-t))\, dt. \tag{4.139}$$

PROOF. Since the vector space $\mathcal{A}_{\mathbb{C}}$ has countable basis and the action of σ_t is unitary, one can find a linear basis of $\mathcal{A}_{\mathbb{C}}$ whose elements are eigenvectors for σ_t and such that the corresponding characters are unitary. Indeed, writing $\mathcal{A}_{\mathbb{C}} = \cup E_n$ with E_n finite dimensional and taking $a \in \mathcal{A}_{\mathbb{C}}$, the set $F_n = \{t \in \mathbb{R} : \sigma_t(a) \in E_n\}$ is a closed subset of $\mathbb{R}$ and $\cup F_n = \mathbb{R}$. It thus follows that for some n the interior of F_n is non-empty so that in fact, using the Taylor expansion at a point inside F_n, the whole orbit $\sigma_t(a)$ is contained in a finite dimensional subspace E_n which allows one to diagonalize the action of σ_t. It follows, using the above linear basis of eigenvectors, that the associative product given by (4.139) leaves $\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ globally invariant. ☐

We denote by $\hat{\mathcal{A}}_{\mathbb{C}}$ the algebra given by $\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ endowed with the product (4.139). At the level of C^*-algebras, one can consider the crossed product $\hat{\mathcal{A}} = \mathcal{A} \rtimes_\sigma \mathbb{R}$ by the one-parameter family of automorphisms $\sigma : \mathbb{R} \to \mathrm{Aut}(\mathcal{A})$ given by the time evolution. The C^*-algebra $\hat{\mathcal{A}}$ is the norm completion of continuous $\mathcal{A}$-valued functions on $\mathbb{R}$ with compact support with the convolution product (4.139). In other words, the elements of the algebra $\hat{\mathcal{A}}$ can be written formally in the form

$$X = \int_{\mathbb{R}} x(t) U_t\, dt, \tag{4.140}$$

where the $\{U_t\}_{t\in\mathbb{R}}$ are unitaries, with product

$$\begin{aligned} \int_{\mathbb{R}} x(t)U_t\, dt \int_{\mathbb{R}} y(r)U_r\, dr &= \int_{\mathbb{R}^2} x(t)\sigma_t(y(r))U_{t+r}\, dt\, dr \\ &= \int_{\mathbb{R}} (x * y)(s)\, U_s\, ds. \end{aligned} \tag{4.141}$$

Thus, the Schwartz space with the product (4.139) gives a dense subalgebra $\hat{\mathcal{A}}_{\mathbb{C}} \subset \hat{\mathcal{A}}$ of the crossed product C^*-algebra.

There is an action of $\mathbb{R}_+^*$ by scaling on the algebra $\hat{\mathcal{A}}$ obtained as follows.

LEMMA 4.59. *There is an action $\theta : \mathbb{R}_+^* \to \mathrm{Aut}(\hat{\mathcal{A}})$ given by*

$$\theta_\lambda\Big(\int x(t)\, U_t\, dt\Big) = \int \lambda^{it}\, x(t)\, U_t\, dt. \tag{4.142}$$

It preserves the subalgebra $\hat{\mathcal{A}}_{\mathbb{C}}$.

PROOF. It is clear by (4.141) and the definition (4.142) that

$$\theta : \mathbb{R}_+^* \to \mathrm{Aut}(\hat{\mathcal{A}})$$

is a continuous one-parameter group of automorphisms that restricts to $\hat{\mathcal{A}}_{\mathbb{C}}$. ☐

Thus, we define the "dual system" in the following way.

DEFINITION 4.60. *The dual system of a quantum statistical mechanical system $(\mathcal{A}, \sigma)$ is the pair $(\hat{\mathcal{A}}, \theta)$ consisting of the algebra $\hat{\mathcal{A}} = \mathcal{A} \rtimes_\sigma \mathbb{R}$ together with the scaling action (4.142). If $\mathcal{A}_{\mathbb{C}}$ is a σ-invariant dense subalgebra of $\mathcal{A}$, then the smooth subalgebra of the dual system is given by $\hat{\mathcal{A}}_{\mathbb{C}}$ as above.*

4.4. Field extensions and duality of factors (an analogy).

We discuss briefly an analogy between passing to the dual system in our context and passing to the algebraic closure $\bar{\mathbb{F}}_q$ in the case of a curve defined over a finite field $\mathbb{F}_q$.

Suppose that C is a smooth projective, geometrically connected, curve over a finite field $\mathbb{F}_q$. Then the function field $\mathbb{K} = \mathbb{F}_q(C)$ is a global field of positive characteristic. It has a family of unramified extensions that come from extensions of the field of coefficients, of the form

$$\mathbb{K} \otimes_{\mathbb{F}_q} \mathbb{F}_{q^n} \quad \text{and} \quad \mathbb{K} \otimes_{\mathbb{F}_q} \bar{\mathbb{F}}_q. \tag{4.143}$$

If $\mathbb{K}$ is instead a number field (a global field of characteristic zero), one does not have a direct analog of these extensions. However, in the characteristic zero case, one finds an analog of these unramified extensions using $\mathcal{A}_{\mathbb{C}} = \mathcal{A}_{\mathbb{K}} \otimes \mathbb{C}$ and then passing to the following crossed products

$$\mathcal{A}_{\mathbb{C}} \rtimes_{\sigma_T} \mathbb{Z} \quad \text{and} \quad \hat{\mathcal{A}}_{\mathbb{C}} = \mathcal{A}_{\mathbb{C}} \rtimes_{\sigma} \mathbb{R}. \tag{4.144}$$

The reason behind this comparison is the following.

Let K be a *local field* with $|\cdot| : K \to \mathbb{R}_+^*$ the absolute value defined by the scaling of the Haar measure on the additive group K induced by the multiplication action of K^* on K. In this way, one can associate to a local field K an invariant $\mathrm{Mod}(K) \subset \mathbb{R}_+^*$ given by

$$\mathrm{Mod}(K) = \{|x| \in \mathbb{R}_+^* \mid x \in K^*\}. \tag{4.145}$$

This is a closed subgroup of $\mathbb{R}_+^*$. The local field is Archimedean iff $\mathrm{Mod}(K) = \mathbb{R}_+^*$ and is non-Archimedean otherwise. In the non-Archimedean case one has

$$\mathrm{Mod}(K) = \lambda^{\mathbb{Z}} \subset \mathbb{R}_+^*, \tag{4.146}$$

for some prime power $\lambda = p^\ell$ for some $\ell \in \mathbb{Z}$, where p^ℓ is the cardinality of the residue field $\mathcal{O}/\mathfrak{m}$ of K. If L is an extension of K of degree n then for an element $x \in K$ the value $|x|_L$ is then the n-th power of $|x|_K$.

In particular, consider a place v of the function field $\mathbb{K} = \mathbb{F}_q(C)$. This defines a corresponding local field $K = \mathbb{K}_v = \mathbb{F}_{q^m}((u))$, where u is a local coordinate (local uniformizer) with $v(u) = 1$. In this case the λ in (4.146) is the cardinality of the residue field $k_v = \mathcal{O}_v/\mathfrak{m}_v = q^{\deg(v)}$.

Let $\mathbb{K}$ be a *global field*. The scaling of the Haar measure on the additive group of adèles $\mathbb{A}_{\mathbb{K}}$, induced by the multiplication action of $\mathbb{A}_{\mathbb{K}}^*$ on $\mathbb{A}_{\mathbb{K}}$, defines the module $|x| \in \mathbb{R}_+^*$ for any $x \in \mathbb{A}_{\mathbb{K}}^*$. The module of the global field $\mathbb{K}$ is the invariant $\mathrm{Mod}(\mathbb{K}) \subset \mathbb{R}_+^*$ given by

$$\mathrm{Mod}(\mathbb{K}) = \{|x| \in \mathbb{R}_+^* \mid x \in \mathbb{A}_{\mathbb{K}}^*\}. \tag{4.147}$$

This is a closed subgroup of $\mathbb{R}_+^*$.

Now, on the side of algebras, we see a similar pattern. Given the data $(\mathcal{A}, \varphi)$ as above one obtains a corresponding von Neumann algebra $\mathcal{M}$ in the GNS representation. We have seen in the previous section the construction of the dual system $\hat{\mathcal{A}} = \mathcal{A} \rtimes_\sigma \mathbb{R}$ at the level of the C^*-algebra and the dense subalgebra $\hat{\mathcal{A}}_{\mathbb{C}}$. Working with von Neumann algebras, one can similarly consider the dual system

$$\hat{\mathcal{M}} = \mathcal{M} \rtimes_{\sigma_t^\varphi} \mathbb{R}. \tag{4.148}$$

It seems that this whole construction depends strongly on the datum of the state φ in the time evolution. However, an important result of Connes' thesis [**63**], (cf. Theorem 4.207 of the Appendix) which is at the basis of the classification of type III factors, shows that in fact the 1-parameter family

$$\sigma^{\varphi} : \mathbb{R} \to \mathrm{Aut}(\mathcal{M}) \tag{4.149}$$

is independent of φ up to inner automorphisms of $\mathcal{M}$. Namely, there is a completely canonical time evolution on a von Neumann algebra $\mathcal{M}$ of the form

$$\delta : \mathbb{R} \to \mathrm{Out}(\mathcal{M}) = \mathrm{Aut}(\mathcal{M})/\mathrm{Inn}(\mathcal{M}), \qquad \delta_t \equiv \sigma_t^{\varphi} \mod \mathrm{Inn}(\mathcal{M}). \tag{4.150}$$

It follows, in particular, that the crossed product (4.148) is *independent* of the choice of φ and one can show (cf. [**99**]) that the dual action θ_λ on $\hat{\mathcal{M}}$ is also independent of the choice of φ.

Moreover, in general for a factor $\mathcal{M}$ one can consider an associated invariant $T(\mathcal{M}) = \mathrm{Ker}(\delta)$, introduced in [**63**], which is the set of periods of the time evolutions σ_t^{φ}. Namely, for any $T \in T(M)$ there exists a faithful normal state φ on $\mathcal{M}$ such that the corresponding time evolution satisfies $\sigma_T^{\varphi} = 1$. For such a *periodic state*, in the construction of the dual system the multiplicative group $\mathbb{R}_+^*$ gets replaced by its discrete subgroup $\lambda^{\mathbb{Z}} \subset \mathbb{R}_+^*$ with $\lambda = \exp(-2\pi/T)$. We shall return to this point later (cf. §7.5) when we discuss some aspects of the adèle class space over a function field $\mathbb{K} = \mathbb{F}_q(C)$, for which the corresponding von Neumann algebra is indeed a factor of type III_λ with $\lambda = 1/q$.

Moreover, another invariant of factors

$$\mathrm{Mod}(\mathcal{M}) \subset \mathbb{R}_+^* \tag{4.151}$$

was introduced in [**63**], as a closed subgroup of $\mathbb{R}_+^*$ defined by the equality

$$S(\mathcal{M}) = \bigcap_{\varphi} \mathrm{Spec}(\Delta_\varphi) \subset \mathbb{R}_+, \tag{4.152}$$

where φ varies among (faithful, normal) states on $\mathcal{M}$, with $\Delta_\varphi = S_\varphi^* S_\varphi$ the corresponding modular operator as in (4.121). The definition of $\mathrm{Mod}(\mathcal{M})$ for arbitrary factors was given in [**99**] (see also [**63**]) as a virtual subgroup of $\mathbb{R}_+^*$ in the sense of Mackey, i.e. an ergodic action of $\mathbb{R}_+^*$. All ergodic flows appear in this way. If we exclude the trivial case $M = M_n(\mathbb{C})$ of matrix algebras, and the fact that $\mathrm{Mod}(\mathcal{M})$ does not distinguish between type II_1 and II_∞, then $\mathrm{Mod}(\mathcal{M})$ is a complete invariant for hyperfinite factors, namely one has $\mathcal{M}_1$ isomorphic to $\mathcal{M}_2$ iff $\mathrm{Mod}(\mathcal{M}_1) \cong \mathrm{Mod}(\mathcal{M}_2)$.

The following statement, which follows from the results of [**63**], shows the analogy with the unramified extensions.

Proposition 4.61. *Let $\mathcal{M}$ be a factor of type III_1. Then the following holds.*

- *For $\lambda \in (0,1)$ let $T = -\frac{2\pi}{\log \lambda}$. Consider the automorphisms $\sigma_T : \mathbb{Z} \to \mathrm{Aut}(\mathcal{M})$ of the form $n \mapsto \sigma_t^{\varphi}(a)$ with $t = nT$. Then the von Neumann algebra*

$$\mathcal{M}_T = \mathcal{M} \rtimes_{\sigma_T} \mathbb{Z} \tag{4.153}$$

is a factor of type III_λ with

$$\mathrm{Mod}(\mathcal{M}_T) = \lambda^{\mathbb{Z}} \subset \mathbb{R}_+^*. \tag{4.154}$$

- *There is a natural inclusion $\mathcal{M} \subset \mathcal{M}_T$ and a dual action*

$$\theta : \mathbb{R}_+^*/\lambda^{\mathbb{Z}} \to \mathrm{Aut}(\mathcal{M}_T) \tag{4.155}$$

of the group $U(1) = \mathbb{R}_+^*/\lambda^{\mathbb{Z}}$, which has $\mathcal{M}$ as fixed points subalgebra

$$\mathcal{M} = \{a \in \mathcal{M}_T \,|\, \theta_\lambda(a) = a, \ \forall \lambda \in \mathbb{R}_+^*/\lambda^{\mathbb{Z}}\}. \tag{4.156}$$

Thus, for a type III_1 factor $\mathcal{M}$ and a given $\lambda \in (0,1)$, we can think of the inclusions $\mathcal{M} \hookrightarrow \mathcal{M}_T$ in the type III_λ factors (4.153) as analogs of the unramified extensions $\mathbb{K} \otimes_{\mathbb{F}_q} \mathbb{F}_{q^n}$ of a function field and of the dual system $\mathcal{M} \rtimes_\sigma \mathbb{R}$ as the analog of the field extension $\mathbb{K} \otimes_{\mathbb{F}_q} \bar{\mathbb{F}}_q$.

We can summarize these observations in the following way:

Function field $\mathbb{K}$	$\mathbf{III}_1$ **Factor** $\mathcal{M}$
$\mathrm{Mod}\,\mathbb{K} \subset \mathbb{R}_+^*$	$\mathrm{Mod}\,\mathcal{M} \subset \mathbb{R}_+^*$
$\mathbb{K} \hookrightarrow \mathbb{K} \otimes_{\mathbb{F}_q} \mathbb{F}_{q^n}$	$\mathcal{M} \hookrightarrow \mathcal{M} \rtimes_{\sigma_T} \mathbb{Z}$
$\mathbb{K} \hookrightarrow \mathbb{K} \otimes_{\mathbb{F}_q} \bar{\mathbb{F}}_q$	$\mathcal{M} \hookrightarrow \mathcal{M} \rtimes_\sigma \mathbb{R}$

Notice that, as observed in Lemma 4.49 above (cf. [**30**]) the von Neumann algebra $\mathcal{M}$ associated to the Bost–Connes endomotive is a type III_1 factor, hence the results of Proposition 4.61 apply. Thus, the analogy illustrated in this section will be especially relevant in the case of the BC system, on which we concentrate later in this chapter.

4.5. Low temperature KMS states and scaling.

As before, consider the data $(\mathcal{A}, \varphi)$ of a unital C^*-algebra with a state φ and a dense subalgebra $\mathcal{A}_{\mathbb{C}}$. Let σ_t^φ be the time evolution obtained via Tomita's theory on the von Neumann algebra $\mathcal{M}$ associated to the GNS representation. We assume as before that σ_t^φ preserves the C^*-algebra and the dense subalgebra. We also assume as before that, for sufficiently large $\beta > 1$, the extremal KMS_β states on $\mathcal{A}$ for the time evolution σ_t^φ are of type I_∞. We then consider the space $\tilde{\Omega}_\beta$ as in Definition 4.55. The following result shows that the points (ε, H) of $\tilde{\Omega}_\beta$ determine irreducible representations of the dual algebra $\hat{\mathcal{A}}$.

Proposition 4.62. *Points $(\varepsilon, H) \in \tilde{\Omega}_\beta$ determine representations of the algebra $\hat{\mathcal{A}} = \mathcal{A} \rtimes_{\sigma^\varphi} \mathbb{R}$ on the Hilbert space $\mathcal{H}_\varepsilon^{(1)}$ of the form*

$$\pi_{\varepsilon,H}\left(\int_{\mathbb{R}} x(t)\, U_t\, dt\right) = \int_{\mathbb{R}} \pi_\varepsilon(x(t))\, e^{itH}\, dt. \tag{4.157}$$

The representations $\pi_{\varepsilon,H}$ are pairwise inequivalent irreducible representations of $\hat{\mathcal{A}}$.

PROOF. It is clear that (4.157) defines a representation of $\hat{\mathcal{A}}$ by bounded operators on the Hilbert space $\mathcal{H}_\varepsilon^{(1)}$. By construction the representations π_ε of $\mathcal{A}$ are already irreducible, so that one obtains that the $\pi_{\varepsilon,H}$ are themselves irreducible. We then need to show that the unitary equivalence class of the representation $\pi_{\varepsilon,H}$ determines the point $(\varepsilon, H) \in \tilde{\Omega}_\beta$. We know that the state ε is uniquely determined by the representation π_ε, cf. Remark 4.54. Consider then a unitary equivalence

$$U^*(\int \pi_\varepsilon(x(t))e^{itH}dt)U = \int \pi_\varepsilon(x(t))e^{it(H+c)}dt.$$

By the irreducibility of the representation π_ε of $\mathcal{A}$ we know that the only operators U that can intertwine the two sides are the scalar operators, which implies $c = 0$. Thus also H is determined uniquely. □

The construction of irreducible representations of $\hat{\mathcal{A}}$ from points of $\tilde{\Omega}_\beta$ is nicely compatible with the scaling action, in the following way.

LEMMA 4.63. *Let $\theta : \mathbb{R}_+^* \to \mathrm{Aut}(\hat{\mathcal{A}})$ be the scaling action* (4.142)*. For all $(\varepsilon, H) \in \tilde{\Omega}_\beta$ this satisfies*

$$\pi_{\varepsilon,H} \circ \theta_\lambda = \pi_{\lambda(\varepsilon,H)}, \qquad \forall \lambda \in \mathbb{R}_+^*, \tag{4.158}$$

with $\lambda(\varepsilon, H)$ the $\mathbb{R}_+^*$-action on $\tilde{\Omega}_\beta$ of (4.136) *and $\pi_{\varepsilon,H}$ the representations* (4.157)*.*

PROOF. We have

$$\pi_{\varepsilon,H}(\theta_\lambda(\int x(t)U_t\,dt)) = \pi_{\varepsilon,H}(\int \lambda^{it}x(t)U_t\,dt)$$
$$= \int \pi_\varepsilon(x(t))\,\lambda^{it}\,e^{itH}\,dt = \int \pi_\varepsilon(x(t))\,e^{it(H+\log\lambda)}\,dt.$$

□

Next, we can show that, for elements of $\hat{\mathcal{A}}_\mathbb{C}$ where the function $x(t) \in \mathcal{S}(\mathbb{R}, \mathcal{A}_\mathbb{C})$ has better regularity properties, the operators $\pi_{(\varepsilon,H)}(\int x(t)U_t\,dt)$ have correspondingly nice spectral properties. This will be very useful in our main application to the case of the BC system.

We first identify a subalgebra of $\hat{\mathcal{A}}_\mathbb{C}$ which depends upon the inverse temperature β. Let $I_\beta = \{z \in \mathbb{C} \,|\, \Im z \in [0, \beta]\}$ denote the same strip in the complex plane that is used to define the KMS$_\beta$ condition. Let $\partial I_\beta = \mathbb{R} \cup \mathbb{R}_\beta$ be its boundary, with $\mathbb{R}_\beta = \{z = t + i\beta \,|\, t \in \mathbb{R}\}$.

PROPOSITION 4.64. *Let $x \in \mathcal{S}(\mathbb{R}, \mathcal{A}_\mathbb{C})$. Then x is the restriction to $\mathbb{R} \subset I_\beta$ of an element of $\mathcal{S}(I_\beta, \mathcal{A}_\mathbb{C})$ if and only if the the following properties hold :*

- *The function $x(t)$ has an analytic continuation $x(z)$ bounded on the strip I_β.*
- *The restriction $x(t + i\beta)$ of $x(z)$ to the boundary is a Schwartz function in $\mathcal{S}(\mathbb{R}_\beta, \mathcal{A}_\mathbb{C})$.*

PROOF. From Definition 4.57 we can assume that x is in fact a scalar-valued function. Then $x \in \mathcal{S}(I_\beta)$ means that all the functions of the form $z^n \frac{\partial}{\partial z}^m x(z)$ are continuous, bounded and holomorphic in the strip I_β. This implies the above two conditions. Conversely, the three line interpolation Theorem [**257**] implies the required boundedness of $z^n \frac{\partial}{\partial z}^m x(z)$ from the assumed bound on the boundary values. □

DEFINITION 4.65. *Let $\hat{\mathcal{A}}_\beta \subset \hat{\mathcal{A}}_\mathbb{C}$ denote the linear space of elements $x \in \mathcal{S}(\mathbb{R}, \mathcal{A}_\mathbb{C})$ with $\mathcal{S}(I_\beta, \mathcal{A}_\mathbb{C})$.*

Definition 4.65 determines a subalgebra of $\hat{\mathcal{A}}_\mathbb{C}$.

LEMMA 4.66. *The space $\hat{\mathcal{A}}_\beta$ defined above is a subalgebra of $\hat{\mathcal{A}}_\mathbb{C}$. It is a norm dense subalgebra of the C^*-algebra $\hat{\mathcal{A}}$.*

PROOF. The space $\mathcal{S}(I_\beta)$ is stable under the convolution product $\int_\mathbb{R} x(u+s)y(u)du$ since taking boundary values commutes with taking the convolution product and the Schwartz space $\mathcal{S}(\mathbb{R})$ is stable under convolution. Using Definition 4.57 and the diagonalization of the action of σ_t one is reduced to the case of eigenvectors and hence of scalar functions, thus proving the first statement.
To show the second it is enough to show that one can approximate any element of $x \in C_c(\mathbb{R}, \mathcal{A})$ by $\mathcal{S}(I_\beta, \mathcal{A}_\mathbb{C})$ in the norm of $L^1(\mathbb{R}, \mathcal{A})$. The norm density of $\mathcal{A}_\mathbb{C}$ in $\mathcal{A}$ and the uniform continuity of x allow one to assume that $x \in C_c(\mathbb{R}, E)$ where $E \subset \mathcal{A}_\mathbb{C}$ is a finite dimensional subspace. The statement then follows from the density of $\mathcal{S}(I_\beta)$ in $L^1(\mathbb{R})$. ☐

Let $\mathcal{L}^1(\mathcal{H})$ denote the ideal of trace class operators in the algebra of bounded operators on a Hilbert space $\mathcal{H}$. We then have the following result.

PROPOSITION 4.67. *For $x \in \hat{\mathcal{A}}_\beta$, one has*

$$\pi_{\varepsilon,H}\Big(\int x(t)\, U_t\, dt\Big) \in \mathcal{L}^1(\mathcal{H}_\varepsilon^{(1)}). \tag{4.159}$$

PROOF. For simplicity of notation, let us consider scalar-valued functions $x \in \mathcal{S}(\mathbb{R}, \mathbb{C})$ at first. In this case, we consider an element of the form

$$\int_\mathbb{R} x(t)\, e^{itH}\, dt \tag{4.160}$$

that plays the role of the element $\int_\mathbb{R} \pi_\varepsilon(x(t))\, e^{itH}\, dt$ in the $\mathcal{A}_\mathbb{C}$-valued case. Let $e(u)$ denote the spectral projections of the Hamiltonian H and let $g(s)$ be the Fourier transform

$$g(s) = \int_\mathbb{R} x(t)\, e^{its}\, dt. \tag{4.161}$$

We can write the element (4.160) in the form

$$\int_\mathbb{R} x(t)\, e^{itH}\, dt = g(H) = \sum_{u \in \mathrm{Spec}(H)} g(u)e(u). \tag{4.162}$$

We assume that the function $x \in \mathcal{S}(I_\beta, \mathbb{C})$. Thus, its Fourier transform g is also in the Schwartz space. The Fourier transform of $x(t + i\beta)$ is $e^{\beta s} g(s)$ and is also in the Schwartz space. Thus, we have obtained

$$e^{\beta s} g(s) \in \mathcal{S}(\mathbb{R}, \mathbb{C}). \tag{4.163}$$

Then, upon writing $\sum_u g(u)e(u) = \sum_u g(u)e^{\beta u}\, e^{-\beta u} e(u)$ one obtains the trace class property, using (4.163) to give a bound on $g(s)\, e^{\beta s}$ independent of $s \in \mathbb{R}$ and then the fact that $\mathrm{Tr}(e^{-\beta H}) < \infty$.

The general case with $x \in \mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ is completely analogous. If $x \in \mathcal{S}(I_\beta, \mathcal{A}_{\mathbb{C}})$ we obtain

$$\pi_{\varepsilon,H}\Big(\int x(t)\, U_t\, dt\Big) = \int \pi_\varepsilon(x(t)) \sum_u e^{itu} e(u)\, dt = \sum_u \pi_\varepsilon(g(u))\, e(u) \tag{4.164}$$

where, as above, the $e(u)$ are the spectral projections of the Hamiltonian H. The same argument as above shows that the function $g(s)e^{\beta s}$ is rapidly decaying so that one obtains an estimate $\|g(s)e^{\beta s}\| < C$ independently of $s \in \mathbb{R}$. Then one can use the fact that $\sum e^{-\beta s}\mathrm{Tr}(e(s)) = \mathrm{Tr}(e^{-\beta H}) < \infty$ to obtain the required trace class property. $\square$

Given a C^*-algebra B we follow [**120**] and let $\mathrm{Prim}\, B$ be the space of primitive ideals of B endowed with the Jacobson topology defined in [**120**] Definition 3.1.5. A primitive ideal is by definition the kernel $\ker \pi$ of an irreducible Hilbert space representation π of B. By [**120**] Theorem 2.9.7 every closed two-sided ideal J of B is the intersection of the primitive ideals containing J. This set of primitive ideals defines a closed subset $\bar{J} \subset \mathrm{Prim}\, B$.

THEOREM 4.68. *For any $(\varepsilon, H) \in \tilde{\Omega}_\beta$ the range of the irreducible representation $\pi_{\varepsilon,H}$ is the algebra of compact operators.*
The map $(\varepsilon, H) \mapsto \ker \pi_{\varepsilon,H}$ gives an $\mathbb{R}_+^*$-equivariant embedding

$$\tilde{\Omega}_\beta \subset \mathrm{Prim}\, \hat{\mathcal{A}}\,. \tag{4.165}$$

PROOF. By Proposition 4.67 and the norm density (Lemma 4.66) of $\hat{\mathcal{A}}_\beta$ in the C^*-algebra $\hat{\mathcal{A}}$, we get that the range of $\pi_{\varepsilon,H}$ is contained in the algebra of compact operators. By [**120**] Corollary 4.1.10, and the irreducibility of $\pi_{\varepsilon,H}$, it follows that the range is the algebra of compact operators. The second statement follows from the second part of [**120**] Corollary 4.1.10. $\square$

We can use the embedding 4.165 and induce the Jacobson topology of $\mathrm{Prim}\, \hat{\mathcal{A}}$ on $\tilde{\Omega}_\beta$. This however does not fully capture the required properties of the corresponding irreducible representations of $\hat{\mathcal{A}}$. This requirement is encoded by imposing the continuity of the functions,

$$(\varepsilon, H) \mapsto \mathrm{Tr}(g(H)) \tag{4.166}$$

for any continuous function g on $\mathbb{R}$ such that $|g(s)e^{\beta s}| < C$ for $s \to +\infty$. We thus endow $\tilde{\Omega}_\beta$ with the weakest topology for which these functions and the embedding (4.165) are continuous. We need to check that this topology is compatible with the action of $\mathbb{R}_+^*$ on $\tilde{\Omega}_\beta$. This follows from the pointwise norm continuity of the scaling action $\theta : \mathbb{R}_+^* \to \mathrm{Aut}(\hat{\mathcal{A}})$ together with Lemma 4.63. The space Ω_β is endowed with the weak topology of the state space of the algebra $\mathcal{A}$.

PROPOSITION 4.69. *The projection $\tilde{\Omega}_\beta \xrightarrow{p} \Omega_\beta$ is continuous.*
For $x \in \hat{\mathcal{A}}_\beta$, the function

$$f(\varepsilon, H) = \mathrm{Tr}\Big(\pi_{\varepsilon,H}\Big(\int x(t)\, U_t\, dt\Big)\Big) \quad \forall (\varepsilon, H) \in \tilde{\Omega}_\beta \tag{4.167}$$

is continuous on $\tilde{\Omega}_\beta$.

PROOF. By Proposition 3.5.9 of [**120**] and the continuity of the embedding (4.165), one gets the lower semicontinuity of the function

$$(\varepsilon, H) \mapsto \mathrm{Tr}(\pi_{\varepsilon,H}(x))$$

for $x \in \hat{\mathcal{A}}$, $x \geq 0$. This implies that the set of $x \in \hat{\mathcal{A}}$, $x \geq 0$ for which this function is continuous forms a face F, and hence defines a two-sided ideal J of $\hat{\mathcal{A}}$ by Lemma 4.5.1 of [**120**], such that $F = J^+$ is the positive part of J.
In particular, the continuity of $(\varepsilon, H) \mapsto \mathrm{Tr}(\pi_{\varepsilon,H}(x))$ holds for any $x \in J$. Using the continuity of the functions (4.166), one gets that J contains the $g(H)$, provided that $g \in C_0(\mathbb{R})$ and $|g(s)e^{\beta s}|$ is bounded for $s \to +\infty$.
Similarly, the continuity of the functions (4.166) gives for $(\varepsilon_0, H_0) \in \tilde{\Omega}_\beta$ an open neighborhood W of (ε_0, H_0) and a constant $c \in \mathbb{R}$ such that $H \geq c$ for all $(\varepsilon, H) \in W$. It is in fact enough to use the continuity of (4.166) for $g(s) = |s| - s$. Let then $g \in C_0(\mathbb{R})$, $g \geq 0$, which agrees with $e^{-\beta s}$ for $s \geq c$. One has $g(H) \in J$ and $xg(H) \in J$, for $x \in \mathcal{A}$.
This gives the continuity of $\mathrm{Tr}(\pi_{\varepsilon,H}(xe^{-\beta H}))$ at the point $(\varepsilon_0, H_0) \in \tilde{\Omega}_\beta$ and proves the continuity of the projection $\tilde{\Omega}_\beta \xrightarrow{p} \Omega_\beta$.
Using the ideal J and the proof of Proposition 4.67 one gets in the same way that for $x \in \hat{\mathcal{A}}_\beta$ one has $x \in J$, which gives the second statement. ☐

4.6. The kernel of the dual trace.

The next results describe the behavior of the function $f(\varepsilon, H)$ of (4.167) on the fibers of the fibration $\mathbb{R}_+^* \to \tilde{\Omega}_\beta \xrightarrow{p} \Omega_\beta$.
We first consider, for simplicity, the case where $x \in \mathcal{S}(I_\beta)$ is a scalar-valued function and we write as above $g(s) = \int_{\mathbb{R}} x(t)\, e^{its}\, dt$ for the Fourier transform. We then have

$$\mathrm{Tr}\left(\pi_{\varepsilon,H+c}\left(\int x(t)\, U_t\, dt\right)\right) = \mathrm{Tr}(g(H+c)). \tag{4.168}$$

We then have the following estimate.

LEMMA 4.70. *Given $x \in \mathcal{S}(I_\beta)$ as above, for all $N > 0$ the following estimate holds:*

$$|\mathrm{Tr}(g(H+c))| = O(e^{-\beta\, c}\, c^{-N}), \qquad \textit{when} \quad c \to +\infty. \tag{4.169}$$

PROOF. We write $g(s) = e^{-\beta\, s} h(s)$ with $h \in \mathcal{S}(\mathbb{R}, \mathbb{C})$. One has

$$|h(u+c)| \leq C\, c^{-N}$$

for $u \in \mathrm{Spec}(H)$, hence $|g(u+c)| \leq C\, e^{-\beta\, c}\, c^{-N}\, e^{-\beta\, u}$, for $u \in \mathrm{Spec}(H)$. This gives

$$|\mathrm{Tr}(g(H+c))| \leq C\, e^{-\beta\, c}\, c^{-N}\, \mathrm{Tr}(e^{-\beta\, H}).$$

☐

We then obtain the following result for the general case of $x \in \hat{\mathcal{A}}_\beta$.

PROPOSITION 4.71. *For any $x \in \hat{\mathcal{A}}_\beta$, the following estimate holds:*

$$|\mathrm{Tr}(\pi_{\varepsilon,H+c}(x))| = O(e^{-\beta\, c}\, c^{-N}), \qquad \textit{for} \quad c \to +\infty. \tag{4.170}$$

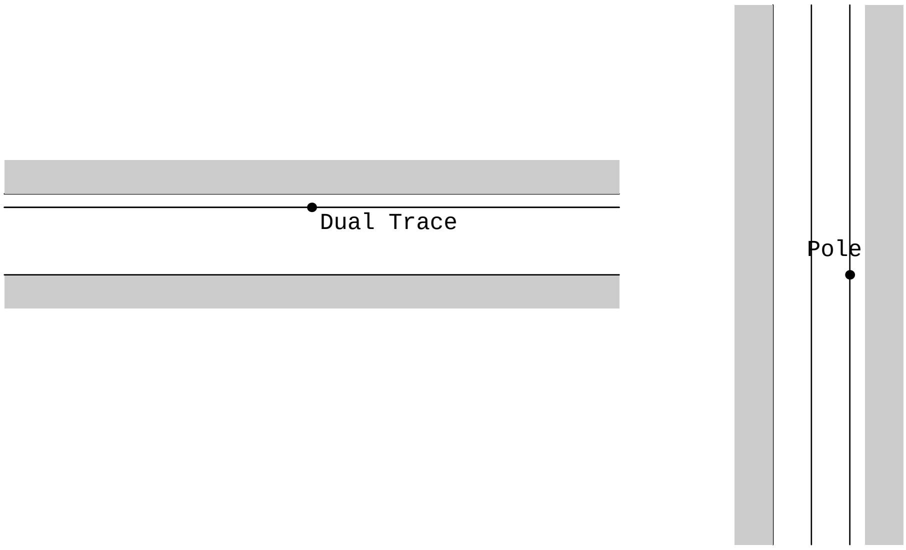


FIGURE 1. The KMS strip and the critical strip

PROOF. We have, as in Proposition 4.67,

$$\mathrm{Tr}(\pi_{\varepsilon,H+c}(x)) = \sum_{u\in\mathrm{Spec}(H)} \mathrm{Tr}(\pi_\varepsilon(g(u+c))\, e(u)).$$

The size of the expression above can then be estimated in terms of

$$\sum_u \|g(u+c)\| \mathrm{Tr}(e(u)).$$

As in Lemma 4.70 above, we estimate

$$\|g(u+c)\| \le C\, e^{-\beta\, c}\, c^{-N}\, e^{-\beta\, u}, \qquad \text{for } u \in \mathrm{Spec}(H).$$

One then obtains the estimate (4.170) using, as in Proposition 4.67, the fact that

$$\sum_{u\in\mathrm{Spec}(H)} \mathrm{Tr}(e(u))e^{-\beta\, u} < \infty.$$

□

Notice that, in general, this does not control the behavior at $c \to -\infty$, due to possible singularities at zero. We see in §4.9 below how one can obtain estimates for the behavior at $c \to -\infty$ in the case of the BC system. In that case one gets

$$\mathrm{Tr}(\pi_{\varepsilon,H+c}(x)) = e^{-c}\,\tau(x) + O(|c|^{-N}), \quad c \to -\infty, \quad \forall N > 0, \tag{4.171}$$

where τ is the canonical *dual trace* (cf. [**280**]) on $\hat{\mathcal{A}}$ which we are now recalling in general. This notion plays an essential role in the theory of factors in the continuous decomposition of type III factors (cf. [**63**], [**280**]).

Let ψ be a KMS$_\beta$ state for a dynamical system $(\mathcal{A}, \sigma_t)$. The dual weight $\hat{\psi}$ on $\hat{\mathcal{A}}$ is defined as

$$\hat{\psi}(\int_{\mathbb{R}} x(t)\, U_t\, dt) = \int_{\mathbb{R}} \psi(\hat{x}(s))\, ds, \tag{4.172}$$

where

$$\hat{x}(s) = \int_{\mathbb{R}} x(t)\, e^{its}\, dt$$

is the Fourier transform with $\hat{x} \in \mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ for $x \in \mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$. If $\beta = 0$ then the dual weight is a trace. For $\beta \neq 0$, and for $X = \int_{\mathbb{R}} x(t)\, U_t\, dt \in \hat{\mathcal{A}}_\beta$, let

$$X_\beta := \int_{\mathbb{R}} x_\beta(t)\, U_t\, dt, \tag{4.173}$$

where $x_\beta(t) = x(t + i\beta)$. The element X_β has the following property.

LEMMA 4.72. *Let $\pi : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ be a covariant representation of $(\mathcal{A}, \sigma_t)$ such that the time evolution has the form $\pi(\sigma_t(a)) = e^{itH} a e^{-itH}$ for a Hamiltonian H. Let $\pi_H : \hat{\mathcal{A}} \to \mathcal{L}(\mathcal{H})$ be the corresponding representation, $\pi_H(X) = \int_{\mathbb{R}} \pi(x(t))\, e^{itH}\, dt$. Then the element X_β as in (4.173) satisfies*

$$\pi_H(X_\beta) = \pi_H(X) e^{\beta H}. \tag{4.174}$$

PROOF. We have

$$\pi_H(X_\beta) = \int_{\mathbb{R}} \pi(x(t + i\beta)) e^{itH}\, dt = \int_{\mathbb{R}} \pi(x(t))\, e^{itH} e^{\beta H}\, dt = \pi_H(X) e^{\beta H}.$$

$\square$

We obtain the dual trace τ_ψ in the following way (cf. [**280**]).

LEMMA 4.73. *Moreover, for ψ a KMS$_\beta$ state for $(\mathcal{A}, \sigma)$, the functional*

$$\tau_\psi(X) := \hat{\psi}(X_\beta) \tag{4.175}$$

is a trace on $\hat{\mathcal{A}}_\beta \subset \mathcal{A} \rtimes_\sigma \mathbb{R}$ satisfying the scaling property

$$\tau_\psi(\theta_\lambda(X)) = \lambda^{-\beta}\, \tau_\psi(X), \tag{4.176}$$

for all $X \in \hat{\mathcal{A}}_\beta$ and all $\lambda \in \mathbb{R}_+^*$.

PROOF. We have by (4.175)

$$\tau_\psi(X) = \int_{\mathbb{R}} \psi(\hat{x}(s))\, e^{\beta s}\, ds. \tag{4.177}$$

The trace property can be seen by writing explicitly

$$\tau_\psi(XY) = \int_{\mathbb{R}} \psi(\widehat{x * y}(s))\, e^{\beta s}\, ds, \tag{4.178}$$

with $x * y$ as in (4.139), and using the KMS$_\beta$ property of the state ψ. The scaling property also follows from (4.177), since we get

$$\tau_\psi(\theta_\lambda X) = \int\int \lambda^{i(t+i\beta)} \psi(x(t + i\beta)) e^{its}\, dt\, ds = \lambda^{-\beta} \tau_\psi(X).$$

$\square$

More generally, one finds that traces on $\hat{\mathcal{A}} = \mathcal{A} \rtimes_\sigma \mathbb{R}$ with the scaling property (4.176) always arise from KMS$_\beta$ states on $(\mathcal{A}, \sigma)$ as above.

DEFINITION 4.74. *For data $(\mathcal{A}, \varphi)$ as above, with $\mathcal{A}_{\mathbb{C}} \subset \mathcal{A}$ a dense subalgebra, let τ_φ denote the trace on $\hat{\mathcal{A}} = \mathcal{A} \rtimes_{\sigma_t^\varphi} \mathbb{R}$ dual to the KMS$_1$ state φ. Let $\hat{\mathcal{A}}^\natural_{\beta,\tau} \subset \hat{\mathcal{A}}^\natural_\beta$ be the cyclic module defined by* $\mathrm{Ker}(\tau_\varphi)$.

The expansion (4.171), in the particular case of the Bost–Connes system, shows why it is important to restrict to the kernel $\hat{\mathcal{A}}^{\natural}_{\beta,\tau}$ of the dual trace τ_φ in order to obtain estimates for the behavior at $c \to -\infty$.

REMARK 4.75. We use the construction of the dual trace for the KMS_1 state φ, but it can also be applied to a KMS_β state $\varepsilon \in \Omega_\beta$. In that case the character of each irreducible representation $\pi_{\varepsilon,H}$ yields a trace on $\hat{\mathcal{A}}_\beta$,

$$\chi_{\varepsilon,H}(x) = \mathrm{Tr}(\pi_{\varepsilon,H}(\int x(t)\, U_t\, dt)) \,. \tag{4.179}$$

The dual trace τ_ε is then given, assuming $\mathrm{Tr}(e^{-\beta H}) = 1$, by

$$\tau_\varepsilon = \int_{\mathbb{R}} \chi_{\varepsilon,H+c}\, e^{\beta c}\, dc \,. \tag{4.180}$$

To get (4.180) one can write, at the formal level,

$$\begin{aligned} \int_{\mathbb{R}} \chi_{\varepsilon,H+c}(x)\, e^{\beta c}\, dc = & \int_{\mathbb{R}} \mathrm{Tr}(\int \pi_\varepsilon(x(t))\, e^{it(H+c)}\, dt)\, e^{\beta c}\, dc \\ = & \; 2\pi\, \mathrm{Tr}(\pi_\varepsilon(x(i\beta))e^{-\beta H}), \end{aligned}$$

while the dual trace τ_ε gives the same result, since one can rewrite (4.172) as

$$\hat{\psi}(\int_{\mathbb{R}} x(t)\, U_t\, dt) = 2\pi\; \psi(x(0)) \tag{4.181}$$

and apply the result for $\psi = \varepsilon = \mathrm{Tr}(\pi_\varepsilon(\cdot)\, e^{-\beta H})$ and $x_\beta(t) = x(t + i\beta)$.

4.7. Holomorphic modules.

We introduce the following algebra of multipliers of $\mathcal{S}(I_\beta)$.

DEFINITION 4.76. *Let $I_\beta = \{z \in \mathbb{C} \,|\, \Im(z) \in [0,\beta]\}$, as before. We let $\mathrm{Hol}(I_\beta)$ denote the algebra of multipliers of $\mathcal{S}(I_\beta)$, i.e. of holomorphic functions h in the strip I_β such that $h\,\mathcal{S}(I_\beta) \subset \mathcal{S}(I_\beta)$.*

For such a function $h \in \mathrm{Hol}(I_\beta)$ the restriction to the boundary component $\mathbb{R}$ of $\partial I_\beta = \mathbb{R} \cup \mathbb{R}_\beta$ is a function of tempered growth and there is a unique distribution $\hat{h}$ on $\mathbb{R}_+^*$ such that

$$\int_{\mathbb{R}_+^*} \hat{h}(\lambda)\lambda^{it}\, d^*\lambda = h(t), \tag{4.182}$$

with $d^*\lambda = \lambda^{-1} d\lambda$ the multiplicative Haar measure on $\mathbb{R}_+^*$.

We then obtain the following general result that refines the result of Lemma 4.63 above.

PROPOSITION 4.77. *Let $\mathrm{Hol}(I_\beta)$ be the algebra of multipliers of $\mathcal{S}(I_\beta)$ as in Definition 4.76.*

(1) *For $h \in \mathrm{Hol}(I_\beta)$, the operator*

$$\theta(\hat{h}) = \int_{\mathbb{R}_+^*} \hat{h}(\lambda)\, \theta_\lambda\, d^*\lambda \tag{4.183}$$

acts on $\hat{\mathcal{A}}_\beta$. This turns $\hat{\mathcal{A}}_\beta$ into a module over the algebra $\mathrm{Hol}(I_\beta)$.

(2) *The map* $\mathbf{F} \circ \mathrm{Tr} \circ \pi$ *is a* $\mathrm{Hol}(I_\beta)$*-module map, where* $\mathbf{F}$ *is the Fourier transform* $\mathbf{F}(g)(u) = \int_{\mathbb{R}} g(c)\, e^{-icu}\, dc$.

PROOF. (1) One has

$$\begin{aligned} \theta(\hat{h})(\textstyle\int x(t)\, U_t\, dt) &= \int_{\mathbb{R}_+^*} \int_{\mathbb{R}} \hat{h}(\lambda)\, \lambda^{it}\, x(t)\, U_t\, dt\, d^*\lambda \\ &= \int_{\mathbb{R}} h(t)\, x(t)\, U_t\, dt, \end{aligned} \tag{4.184}$$

where $hx \in \mathcal{S}(I_\beta, \mathcal{A}_{\mathbb{C}})$.
(2) The equivariance of π proved in Lemma 4.63 namely the fact that

$$\pi_{\varepsilon,H} \circ \theta_\lambda = \pi_{(\varepsilon,H+\log\lambda)}, \qquad \forall (\varepsilon, H) \in \tilde{\Omega}_\beta, \tag{4.185}$$

gives

$$(\mathrm{Tr} \circ \pi \circ \theta_\lambda)(X)_{(\varepsilon,H)} = (\mathrm{Tr} \circ \pi)(X)_{(\varepsilon,H+\log\lambda)}. \tag{4.186}$$

Thus, for the function

$$f_\varepsilon(c) = \mathrm{Tr}(\pi_{\varepsilon,H+c}(X)) \tag{4.187}$$

one obtains

$$(\mathrm{Tr} \circ \pi \circ \theta(\hat{h}))(X)_{(\varepsilon,H)} = \int_{\mathbb{R}_+^*} \hat{h}(\lambda)\, f_\varepsilon(\log\lambda)\, d^*\lambda. \tag{4.188}$$

Replacing H by $H + c$ one gets that the function $g_\varepsilon(c)$ associated to $\theta(\hat{h})(x)$ in the same way as in (4.187) is given by

$$g_\varepsilon(c) = \int_{\mathbb{R}_+^*} \hat{h}(\lambda)\, f_\varepsilon(\log\lambda + c)\, d^*\lambda. \tag{4.189}$$

One then obtains the Fourier transform

$$\begin{aligned} \mathbf{F}(g_\varepsilon)(u) &= \int_{\mathbb{R}} g_\varepsilon(c)\, e^{-icu}\, dc \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}_+^*} \hat{h}(\lambda)\, f_\varepsilon(\log\lambda + c)\, e^{-icu}\, dc\, d^*\lambda \\ &= \int_{\mathbb{R}_+^*} \hat{h}(\lambda)\, \lambda^{iu}\, d^*\lambda \int_{\mathbb{R}} f_\varepsilon(a)\, e^{-iau}\, da \\ &= h(u)\mathbf{F}(f_\varepsilon)(u). \end{aligned} \tag{4.190}$$

□

4.8. The cooling morphism (distillation).

We have described in §4.5 the procedure that associates to the data $(\mathcal{A}, \varphi)$ the space $\tilde{\Omega}_\beta$ and the dual system $(\hat{\mathcal{A}}, \theta_\lambda)$. We shall now explain how this gives rise to a morphism of Λ-modules, whose cokernel in the abelian category of Λ-modules defines a cohomology with an $\mathbb{R}_+^*$ action. This procedure, called *distillation* ([**74**]), will be mainly illustrated in the case of the BC system in §4.9 and we shall give only a brief description of the general case in the present section.
In the general case one is dealing with the field of elementary C^*-algebras

$$C_{\varepsilon,H} = \pi_{\varepsilon,H}(\hat{\mathcal{A}})\,,\ (\varepsilon, H) \in \tilde{\Omega}_\beta \tag{4.191}$$

on $\tilde{\Omega}_\beta$. By Theorem 4.68 the algebras $C_{\varepsilon,H}$ are isomorphic to the algebra of compact operators in a Hilbert space. The Dixmier–Douady invariant (cf. [**120**] Theorem 10.8.4) gives, in general, an obstruction to the Morita equivalence between the algebra $C(\tilde{\Omega}_\beta, C)$ of continuous sections of the field C of elementary C^*-algebras $C_{\varepsilon,H}$ and the algebra of continuous functions on $\tilde{\Omega}_\beta$. We denote by $\mathcal{L}^1$ the functor defined on elementary C^*-algebras by

$$\mathcal{L}^1(\text{compact operators in } \mathcal{H}) = \text{trace class operators in } \mathcal{H}\,. \tag{4.192}$$

Definition 4.78. *The cooling morphism is the canonical map*

$$\begin{aligned} \pi : \hat{\mathcal{A}}_\beta &\to C(\tilde{\Omega}_\beta, \mathcal{L}^1(C)) \\ \pi(x)(\varepsilon, H) &= \tilde{\pi}_{\varepsilon,H}(x), \qquad \forall (\varepsilon, H) \in \tilde{\Omega}_\beta. \end{aligned} \tag{4.193}$$

which defines an $\mathbb{R}_+^*$-equivariant algebra homomorphism.

Proposition 4.69 ensures the continuity of the trace $\mathrm{Tr}(\pi(x)(\varepsilon, H))$ on $\tilde{\Omega}_\beta$.

We then consider the cokernel of the cooling map as an object in the category of Λ-modules. Let $\pi^\natural$ be the morphism of cyclic modules associated to (4.193). The *distilled* Λ-module associated to the data $(\mathcal{A}_\mathbb{C}, \varphi)$ is defined as

$$D_\tau(\mathcal{A}_\mathbb{C}, \varphi) = \mathrm{Coker}(\pi^\natural)\,, \tag{4.194}$$

where the domain of $\pi^\natural$ is the cyclic module $\hat{\mathcal{A}}^\natural_{\beta,\tau}$ defined by the kernel of the dual trace in Definition 4.74. We shall not discuss the exact definition of the range, but we shall see below in the case of the BC system the relevant decay condition to be imposed on the function $\mathrm{Tr}(\pi(x)(\varepsilon, H))$.

We then consider a corresponding cohomology theory provided by the cyclic homology

$$HC_*(D_\tau(\mathcal{A}_\mathbb{C}, \varphi))$$

of the distilled Λ-module. In the following, since we are mostly interested in the case of data $(\mathcal{A}, \varphi)$ that come from endomotives, which are zero-dimensional noncommutative spaces, we will restrict to considering only $HC_0(D_\tau(\mathcal{A}_\mathbb{C}, \varphi)$.

We shall now assume that the Dixmier–Douady invariant is trivial and in fact we shall simplify even further by assuming that the corresponding bundle of Hilbert spaces is trivial. This holds, for instance, in the particular case of the Bost–Connes system. In this case one has $\Omega_\beta = \mathcal{E}_\beta$ (all extremal KMS_β states with $\beta > 1$ are regular) and the representations $\pi_{\varepsilon,H}$ are all defined on the same Hilbert space $\mathcal{H}^{(1)}_\varepsilon = \ell^2(\mathbb{N})$.

We fix a separable Hilbert space $\mathcal{H}$ and let, as in [**120**], $\mathrm{Irr}\,\hat{\mathcal{A}}$ be the space of irreducible representations of the C^*-algebra $\hat{\mathcal{A}}$ in $\mathcal{H}$ endowed with the topology of pointwise weak convergence. Thus

$$\pi_\alpha \to \pi \iff \langle \xi, \pi_\alpha(x)\eta\rangle \to \langle \xi, \pi(x)\eta\rangle\,, \quad \forall\, \xi, \eta \in \mathcal{H},\, x \in \hat{\mathcal{A}}\,. \tag{4.195}$$

This is equivalent to pointwise strong convergence by [**120**] 3.5.2. By [**120**] Theorem 3.5.8 the map

$$\pi \in \mathrm{Irr}\hat{\mathcal{A}} \mapsto \ker \pi \in \mathrm{Prim}\,\hat{\mathcal{A}} \tag{4.196}$$

is continuous and open. Indeed, by Definition 3.1.5 of [**120**], the topology on the spectrum, i.e. the space of equivalence classes of irreducible representations, is *defined* as the pullback of the Jacobson topology of the primitive ideal space.
We assume that the representations $\pi_{\varepsilon,H}$, for $(\varepsilon,H)\in\tilde{\Omega}_\beta$ described in §4.5 can be continuously realized as representations $\tilde{\pi}_{\varepsilon,H}$ in $\mathcal{H}$, i.e. more precisely that the injective continuous map (4.165) is lifted to a continuous map to $\operatorname{Irr}\hat{\mathcal{A}}$ such that

$$(\varepsilon,H)\mapsto\tilde{\pi}_{\varepsilon,H}(\hat{f}(H)) \tag{4.197}$$

is a continuous map from $\tilde{\Omega}_\beta$ to $\mathcal{L}^1=\mathcal{L}^1(\mathcal{H})$ for any $f\in\mathcal{S}(I_\beta)$. This last condition corresponds to (4.166) in the definition of the topology of $\tilde{\Omega}_\beta$. We then show that these representations assemble to give an $\mathbb{R}_+^*$-equivariant morphism.

LEMMA 4.79. *Under the above hypothesis the map*

$$\begin{gathered}\pi:\hat{\mathcal{A}}_\beta\to C(\tilde{\Omega}_\beta,\mathcal{L}^1)\\ \pi(x)(\varepsilon,H)=\tilde{\pi}_{\varepsilon,H}(x),\qquad\forall(\varepsilon,H)\in\tilde{\Omega}_\beta.\end{gathered} \tag{4.198}$$

defines an $\mathbb{R}_+^*$-equivariant algebra homomorphism.

PROOF. It is enough to show that for any $a\in\mathcal{A}_{\mathbb{C}}$ and $f\in\mathcal{S}(I_\beta)$ the map

$$(\varepsilon,H)\mapsto\tilde{\pi}_{\varepsilon,H}(a\,\hat{f}(H)) \tag{4.199}$$

is continuous from $\tilde{\Omega}_\beta$ to $\mathcal{L}^1(\mathcal{H})$. The general result follows since linear combinations of these elements spans $\hat{\mathcal{A}}_\beta$. One can assume that $f=f_1\star f_2$ with $f_1\in L^1(\mathbb{R})$, $f_2\in\mathcal{S}(I_\beta)$. The map $(\varepsilon,H)\mapsto\tilde{\pi}_{\varepsilon,H}(a\,\hat{f}_1(H))$ is bounded and pointwise strongly continuous by (4.195). Thus by (4.197) we get the continuity of (4.199). The equivariance follows from Lemma 4.63. ☐

The morphism (4.198) agrees with (4.193). One can compose it with the trace on $\mathcal{L}^1(\mathcal{H})$ and obtain a morphism in the category of Λ-modules (cf. §1.2 above, and especially §1.3).

LEMMA 4.80. *The morphism* (4.198) *induces a cyclic morphism*

$$\delta\,:\,\hat{\mathcal{A}}^\natural_{\beta,\tau}\overset{(\mathrm{Tr}\circ\pi)^\natural}{\longrightarrow}C(\tilde{\Omega}_\beta)^\natural. \tag{4.200}$$

PROOF. The result follows by applying the formulas of Proposition 4.10 to define $\mathrm{Tr}^\natural$ which is then composed with $\pi^\natural$ for π the cooling morphism (4.198). ☐

Notice that, in this case, because the algebra $C(\tilde{\Omega}_\beta)$ is commutative, the cyclic homology $HC_0(D_\tau(\mathcal{A},\varphi))$ is simply given by the cokernel of δ in degree 0. The advantage of describing this cokernel as an HC_0 lies in the Morita invariance of cyclic homology which, for instance, wipes out the distinction between an algebra and its tensor product with $\mathcal{L}^1$. This cyclic homology $HC_0(D_\tau(\mathcal{A}_{\mathbb{C}},\varphi)$ comes endowed with an induced representation of $\mathbb{R}_+^*$.

LEMMA 4.81. *The cyclic homology $HC_0(D_\tau(\mathcal{A}_{\mathbb{C}},\varphi))$ carries an induced action*

$$\theta:\mathbb{R}_+^*\to\operatorname{Aut}(HC_0(D_\tau(\mathcal{A}_{\mathbb{C}},\varphi))). \tag{4.201}$$

PROOF. The result is a direct consequence of the $\mathbb{R}_+^*$-equivariance of the morphism $\delta=(\mathrm{Tr}\circ\pi)^\natural$, which follows from Lemma 4.79. ☐

The spectrum of this $\mathbb{R}_+^*$ representation on the cyclic homology of $D(\mathcal{A}_{\mathbb{C}}, \varphi)$ defines a very subtle invariant of the original noncommutative space $(\mathcal{A}, \varphi)$, as we are going to see in detail in the Bost–Connes case.

4.9. Distillation of the Bost–Connes endomotive.

In this section we concentrate on the case of the Bost–Connes endomotive, applying the cooling and distillation procedures described in the previous sections. The main result is that the map $\mathfrak{E}$ discussed in Chapter 2, which plays a crucial role in the spectral realization of zeros of the Riemann zeta function and in the (semi-local) trace formula of [**71**], acquires now a cohomological description as the cooling map (4.200) for the BC endomotive.

We first return to the geometric description of the BC system in terms of commensurability classes of $\mathbb{Q}$-lattices, as discussed at length in Chapter 3 (cf. [**86**]). Recall that we denoted by $\mathcal{G}_1$ the groupoid of the equivalence relation of commensurability on 1-dimensional $\mathbb{Q}$-lattices, and by $\mathcal{G}_1/\mathbb{R}_+^*$ the quotient by the scaling action. By Lemma 3.21 of Chapter 3, $\mathcal{G}_1$ is the étale groupoid of triples $(k, \rho, \lambda) \in \mathbb{Q}_+^* \times \hat{\mathbb{Z}} \times \mathbb{R}_+^*$ such that $k\rho \in \hat{\mathbb{Z}}$ with the composition (3.34). By Proposition 3.22 of Chapter 3, $\mathcal{G}_1/\mathbb{R}_+^*$ is the étale groupoid of pairs $(k, \rho) \in \mathbb{Q}_+^* \times \hat{\mathbb{Z}}$ such that $k\rho \in \hat{\mathbb{Z}}$ with the composition (3.37). The algebra $\mathcal{A}_1$ of the BC system $(\mathcal{A}_1, \sigma_t)$ is the convolution algebra

$$\mathcal{A}_1 = C^*(\mathcal{G}_1/\mathbb{R}_+^*), \tag{4.202}$$

which we have identified in §2.7 above with the C^*-algebra of the analytic endomotive obtained from the self-maps of $(\mathbb{G}_m, 1)$. We also consider here the algebra $C^*(\mathcal{G}_1)$ describing the noncommutative space of commensurability classes of 1-dimensional $\mathbb{Q}$-lattices (not up to scale). This is related to the dual system $\hat{\mathcal{A}}_1$ in the following way.

LEMMA 4.82. *There is a C^*-algebra isomorphism $\iota : \hat{\mathcal{A}}_1 \to C^*(\mathcal{G}_1)$ of the form*

$$\iota(X)(k, \rho, \lambda) = \int_{\mathbb{R}} x(t)(k, \rho)\, \lambda^{it}\, dt \tag{4.203}$$

for $(k, \rho, \lambda) \in \mathcal{G}_1$ and

$$X = \int x(t)\, U_t\, dt \in \hat{\mathcal{A}}_1.$$

PROOF. It suffices to check the compatibility of the convolution product (4.139) with the convolution product in the groupoid algebra $C^*(\mathcal{G}_1)$. We have

$$\iota(X_1 X_2)(k, \rho, \lambda) = \int\int (x_1(t) * \sigma_t(x_2(s-t)))(k, \rho) \lambda^{is} dt ds$$

$$= \sum_{r, r\rho \in \hat{\mathbb{Z}}} \int\int x_1(t)(kr^{-1}, r\rho) \sigma_t(x_2(s-t))(r, \rho) \lambda^{is} dt ds$$

$$= \sum_{r, r\rho \in \hat{\mathbb{Z}}} \int\int x_1(t)(kr^{-1}, r\rho) x_2(s-t)(r, \rho) r^{it} \lambda^{is} dt ds$$

$$= \sum_{r, r\rho \in \hat{\mathbb{Z}}} \int\int x_1(t)(kr^{-1}, r\rho) x_2(u)(r, \rho) r^{it} \lambda^{i(t+u)} dt du$$

$$= \sum_{r, r\rho \in \hat{\mathbb{Z}}} \iota(X_1)(kr^{-1}, r\rho, r\lambda)\iota(X_2)(r, \rho, \lambda)\,.$$

The compatibility of the involutions follows similarly. ☐

We also have the following result.

LEMMA 4.83. *Let $\mathcal{A}$ be the algebra of the BC system and $\iota : \hat{\mathcal{A}} \to C^*(\mathcal{G}_1)$ be the homomorphism of (4.203). The range under ι of the subalgebra $\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ consists of functions $f(k, \rho, \lambda)$ with finite support in $k \in \mathbb{Q}_+^*$, finite level in $\rho \in \hat{\mathbb{Z}}$ and rapid decay (Schwartz functions) in $\lambda \in \mathbb{R}_+^*$.*

PROOF. The result follows from the definition of the subalgebra $\mathcal{S}(\mathbb{R}, \mathcal{A}_{\mathbb{C}})$ of $\hat{\mathcal{A}}$ as the direct limit (4.138) over finite dimensional subspaces of $\mathcal{A}_{\mathbb{C}}$ and the fact that the dense subalgebra $\mathcal{A}_{\mathbb{C}}$ of the BC C^*-algebra $\mathcal{A}_1$ is given by the algebraic crossed product of the algebra of locally constant functions on $\hat{\mathbb{Z}}$ by the partial action of $\mathbb{Q}_+^*$. ☐

We can further refine the result of Lemma 4.83 above. As in Chapter 2 (2.143) we use the bicharacter

$$\mu(v, s) = v^{-is}\,, \quad \forall v \in \mathbb{R}_+^*, s \in \mathbb{R}\,, \tag{4.204}$$

and let $\mathbf{F}_\mu$ be the Fourier transform $\mathbf{F}_\mu : L^2(\mathbb{R}_+^*) \to L^2(\mathbb{R})$ associated to μ,

$$\mathbf{F}_\mu(f)(s) := \int_0^\infty f(v) v^{-is} d^*v\,. \tag{4.205}$$

where d^*v is the multiplicative Haar measure on $\mathbb{R}_+^*$.

LEMMA 4.84. *Let $\mathcal{A}$ be the algebra of the BC system and $\iota : \hat{\mathcal{A}} \to C^*(\mathcal{G}_1)$ be the homomorphism of (4.203). The range under ι of the subalgebra $\hat{\mathcal{A}}_\beta$ consists of functions $f(k, \rho, \lambda)$ with finite support in k, finite level in ρ and such that the finitely many scalar functions*

$$f_{k,\rho}(\lambda) = f(k, \rho, \lambda), \qquad \textit{for } \lambda \in \mathbb{R}_+^* \tag{4.206}$$

satisfy $\mathbf{F}_\mu(f_{k,\rho})(u) \in \mathcal{S}(I_\beta)$ for all (k, ρ).

PROOF. The Fourier transform $\mathbf{F}_\mu$ in the variable $\lambda \in \mathbb{R}_+^*$ corresponds to the usual Fourier transform in the variable $s = \log(\lambda)$, so that we have

$$\begin{aligned}
\mathbf{F}_\mu(f_{k,\rho})(u) &= \int_{\mathbb{R}_+^*} \int_{\mathbb{R}} x(t)(k, \rho)\, \lambda^{it}\, dt\, \lambda^{-iu} \frac{d\lambda}{\lambda} \\
&= \int_{\mathbb{R}\times\mathbb{R}} x(t)(k, \rho)\, e^{its}\, e^{-ius}\, dt\, ds \\
&= 2\pi x(u)(k, \rho)
\end{aligned}$$

Then, by Proposition 4.64, the condition $\mathbf{F}_\mu(f_{k,\rho}) \in \mathcal{S}(I_\beta)$ is equivalent to $X = \int x(t) U_t\, dt \in \hat{\mathcal{A}}_\beta$. This gives the characterization

$$f \in \iota\hat{\mathcal{A}}_\beta \Longleftrightarrow \mathbf{F}_\mu(f_{k,\rho}) \in \mathcal{S}(I_\beta), \quad \forall (k, \rho). \tag{4.207}$$

☐

The groupoid $\mathcal{G}_1$ describes the partial action of $\mathbb{Q}_+^*$ on $\hat{\mathbb{Z}}\times\mathbb{R}_+^*$ or equivalently of $\mathbb{Q}^*$ on $\hat{\mathbb{Z}}\times\mathbb{R}^*$ (cf. the description in terms of the Shimura variety of $\mathbb{G}_m$ in Chapter 3). We have seen that the quotient space of the commensurability relation on 1-dimensional $\mathbb{Q}$-lattices is described in this way by the quotient $\mathbb{A}^{\cdot}_{\mathbb{Q}}/\mathbb{Q}^*$, with $\mathbb{A}^{\cdot}_{\mathbb{Q}} = \mathbb{A}_{\mathbb{Q},f}\times\mathbb{R}^*$, where we describe the quotient as a noncommutative space via the convolution algebra $C^*(\mathcal{G}_1)$. In fact, we can by a Morita equivalence pass from the quotient $\mathbb{A}^{\cdot}_{\mathbb{Q}}/\mathbb{Q}^*$, which corresponds to the algebra $C_0(\mathbb{A}^{\cdot}_{\mathbb{Q}})\rtimes\mathbb{Q}^*$ to the algebra $C^*(\mathcal{G}_1)$, where the Morita equivalence comes from the strict inclusion $\hat{\mathbb{Z}}\subset\mathbb{A}_{\mathbb{Q},f}$ of $\hat{\mathbb{Z}}$ in finite adèles. The space $\mathbb{A}^{\cdot}_{\mathbb{Q}}/\mathbb{Q}^*$ is however not the same as the adèle class space $\mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$ that we encountered in the discussion of the Riemann zeta function in Chapter 2, Definition 2.45. Indeed, with respect to the adèle class space $\mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^*$ we are still missing the compact piece $\hat{\mathbb{Z}}\times\{0\}\subset\mathbb{A}_{\mathbb{Q}}$. Still, even though we are now working with the space $\mathbb{A}^{\cdot}_{\mathbb{Q}}/\mathbb{Q}^*$, we can give a description of the map $\mathfrak{E}$ that we saw in Chapter 2, §6 and §9, in terms of the BC endomotive.

Recall that we described the spectral realization of the critical zeros in Chapter 2 §9 from the cokernel of the following map,

$$\mathfrak{E}(f)(g) = |g|^{1/2}\sum_{q\in\mathbb{Q}^*} f(qg),\quad \forall g\in C_{\mathbb{Q}} = \mathbb{A}_{\mathbb{Q}}^*/\mathbb{Q}^*. \tag{4.208}$$

which is equivariant in the sense of (2.357) with respect to the natural scaling actions of $C_{\mathbb{Q}}$, as we discussed in detail in Chapter 2 above.

The following result recovers the map $\mathfrak{E}$ in terms of the cooling map for the BC endomotive.

THEOREM 4.85. *Let $(\mathcal{A},\sigma)$ be the C^*-dynamical system of the Bost–Connes analytic endomotive.*

(1) *For $\beta>1$ there is a canonical isomorphism*

$$\tilde{\Omega}_\beta \simeq \hat{\mathbb{Z}}^*\times\mathbb{R}_+^*\simeq C_{\mathbb{Q}} \tag{4.209}$$

of $\tilde{\Omega}_\beta$ with the space of invertible 1-dimensional $\mathbb{Q}$-lattices.

(2) *For $X\in\hat{\mathcal{A}}$ and $f=\iota(X)\in C^*(\mathcal{G}_1)$, the cooling map* (4.200) *takes the form*

$$\delta(X)(u,\lambda) = \sum_{n\in\mathbb{N}=\mathbb{Z}_{>0}} f(1,nu,n\lambda),\quad \forall (u,\lambda)\in C_{\mathbb{Q}}\simeq\tilde{\Omega}_\beta. \tag{4.210}$$

PROOF. (1) We give explicitly the covariant irreducible representation of $\mathcal{A}$ associated to a pair $(u,\lambda)\in\hat{\mathbb{Z}}^*\times\mathbb{R}_+^* = C_{\mathbb{Q}}$. The algebra acts on the fixed Hilbert space $\mathcal{H}=\ell^2(\mathbb{N})$ as

$$(\pi_{(u,\lambda)}(x)\xi)(n) = \sum_{m\in\mathbb{N}} x(nm^{-1},mu)\xi(m),\qquad \forall x\in\mathcal{A},\ \forall\xi\in\ell^2(\mathbb{N}), \tag{4.211}$$

while $\pi_{(u,\lambda)}(H)$ is the diagonal operator D_λ on $\ell^2(\mathbb{N})$ with

$$(D_\lambda\,\xi)(n) = (\log n+\log\lambda)\xi(n)\,,\ \forall\xi\in\ell^2(\mathbb{N}), \tag{4.212}$$

so that one has

$$\mathrm{Spec}(D_\lambda) = \{\log n+\log\lambda\}_{n\in\mathbb{N}}. \tag{4.213}$$

The classification of KMS$_\beta$ states of the BC system ([**30**], [**86**], see Chapter 3 above) shows that the map $(u,\lambda) \to \pi_{(u,\lambda)}$ gives an isomorphism of $C_{\mathbb{Q}} = \hat{\mathbb{Z}}^* \times \mathbb{R}_+^*$ with $\tilde{\Omega}_\beta$, for all $\beta > 1$.

(2) By construction the function $\delta(x)(u,\lambda)$, with $(u,\lambda) \in C_{\mathbb{Q}}$, is given as in (4.200) by

$$\delta(x)(u,\lambda) = \mathrm{Tr}(\pi_{(u,\lambda)}(x)). \tag{4.214}$$

For $X = \int x(t) U_t \, dt \in \hat{\mathcal{A}}$ one has

$$\pi_{(u,\lambda)}(X) = \int \pi_{(u,\lambda)}(x(t)) \, e^{itD_\lambda} \, dt\,. \tag{4.215}$$

Using the explicit description of the representation given in (4.211) we get the trace as the sum of the diagonal entries of this matrix, which gives

$$\sum_{n\in\mathbb{N}} \int x(t)(1,nu) \, e^{it(\log n + \log \lambda)} \, dt = \sum_{n\in\mathbb{N}} f(1,nu,n\lambda), \tag{4.216}$$

for $f = \iota(X) \in C^*(\mathcal{G}_1)$. ☐

One can compare directly the right hand side of (4.216) with the map $\mathfrak{E}$ (up to the normalization by $|j|^{1/2}$) written as in (4.208) by considering a function $f(\rho,v) = f(1,\rho,v)$ and its unique extension $\tilde{f}$ to adèles obtained by setting $\tilde{f}$ equal to zero outside of $\hat{\mathbb{Z}} \times \mathbb{R}^+ \subset \mathbb{A}_{\mathbb{Q}}$. This then gives

$$\sum_{n\in\mathbb{N}} f(1,nu,n\lambda) = \sum_{q\in\mathbb{Q}^*} \tilde{f}(q\,j), \qquad \text{where } j = (u,\lambda) \in C_{\mathbb{Q}}. \tag{4.217}$$

Notice that, unlike in the case of [**71**], here we do not work with Hilbert spaces but with spaces of rapidly decaying functions. A reformulation of the results of [**71**] in the analytic setting of algebras of rapidly decaying functions and nuclear spaces was given by Ralf Meyer in [**224**]. Our cohomological formulation here uses the same analytic tools as [**224**].

In particular, in the case of the BC system, we can use the result of Theorem 4.85 above to improve the decay estimate of Proposition 4.71 in the following way.

Let us return to the question of the behavior as $c \to -\infty$ of $\mathrm{Tr}(\pi_{\varepsilon,H+c}(X))$, after the general estimate for $c \to +\infty$ given in Proposition 4.71 above. As above, we write $\mathrm{Tr}(\pi_{\varepsilon,H+c}(X)) = \mathrm{Tr}(g(H+c))$ for $X = \int x(t)\, U_t \, dt$ and $g(s) = \int_{\mathbb{R}} x(t)\, e^{its} \, dt$. In the case of the BC system, we have

$$\mathrm{Tr}(g(H+c)) = \sum_{n\in\mathbb{N}} g(\log(n) + c), \tag{4.218}$$

where $\mathrm{Spec}(H) = \{\log(n) \,|\, n \in \mathbb{N} = \mathbb{Z}_{>0}\}$.

Lemma 4.86. *With the notation as above, let $\ell(t) = g(\log(t))$ be extended to $\mathbb{R}$ as a continuous even function. Then for $c \to -\infty$ (i.e. for $\lambda = e^c \to 0$) the leading behavior of* $\mathrm{Tr}(g(H+c))$ *is given by*

$$\sum_{n\in\mathbb{Z}} \ell(\lambda n) \sim \lambda^{-1} \int \ell(t)\, dt, \qquad \textit{for } \lambda \to 0. \tag{4.219}$$

PROOF. The function $\ell(t) = g(\log(t))$ extended to $\mathbb{R}$ has $\ell(0) = 0$ and, for $\lambda = e^c$, one can write from (4.218)

$$\mathrm{Tr}(g(H+c)) = \frac{1}{2}\sum_{n\in\mathbb{Z}} \ell(\lambda n). \tag{4.220}$$

With the appropriate normalization of the Fourier transform $\hat{\ell}$ of ℓ, the Poisson summation formula then gives

$$\lambda \sum_{n\in\mathbb{Z}} \ell(\lambda\, n) = \sum_{n\in\mathbb{Z}} \hat{\ell}(\lambda^{-1}\, n). \tag{4.221}$$

This gives the desired estimate (4.219). ☐

REMARK 4.87. Notice how the remainder is controlled by the decay of the Fourier transform $\hat{\ell}$ at ∞, i.e. by the smoothness of the function $\ell(t) = g(\log(t))$. Notice that the latter is not guaranteed by that of g itself, because of a possible singularity at zero.

In fact, more precisely, we obtain the following result on the behavior at $c \to -\infty$, which gives more precise information on the divergent term and the remainder, and relates it to the kernel of the dual trace discussed in §4.6 above.

We first compute explicitly the dual trace for the Bost–Connes system.

LEMMA 4.88. *Let $(\mathcal{A}, \varphi)$ be the data of the Bost–Connes measured endomotive, with the state φ as in* (4.126). *The dual trace on $\hat{\mathcal{A}} = \mathcal{A} \rtimes_{\sigma_t^\varphi} \mathbb{R}$ is of the form*

$$\tau_\varphi(X) = \int_{\hat{\mathbb{Z}}\times\mathbb{R}_+^*} f(1,\rho,\lambda)\, d\mu(\rho)\, d\lambda, \tag{4.222}$$

for $f = \iota(X)$ under the map (4.203), *with $X = \int x(t) U_t\, dt \in \hat{\mathcal{A}}$. Here $d\mu(\rho)$ is the Haar measure on $\hat{\mathbb{Z}}$ as in* (4.126).

PROOF. By the general description (4.177) of the traces on $\hat{\mathcal{A}}$ dual to KMS states on $(\mathcal{A}, \sigma_t)$ the trace τ_φ is of the form

$$\tau_\varphi(X) = \int_{\mathbb{R}\times\mathbb{R}} \varphi(x(t)) e^{its}\, e^s\, dt\, ds. \tag{4.223}$$

We rewrite this in terms of $\lambda = e^s \in \mathbb{R}_+^*$ as

$$\int_{\mathbb{R}\times\mathbb{R}_+^*} \left(\int_{\hat{\mathbb{Z}}} x(t)(1,\rho)\, d\mu(\rho) \right) \lambda^{it}\, dt\, d\lambda =$$

$$\int_{\hat{\mathbb{Z}}\times\mathbb{R}_+^*} \left(\int_{\mathbb{R}} x(t)(1,\rho)\, \lambda^{it}\, dt \right) d\mu(\rho)\, d\lambda.$$

The right hand side agrees with the expression (4.222) by (4.203). ☐

We then have the following preliminary result.

PROPOSITION 4.89. *Consider a function $f(\rho,v)=f(1,\rho,v)$ such that*

$$f(\rho,v)=g(p_n(\rho),v)$$

with g a map from $\mathbb{Z}/n\mathbb{Z}$ to $\mathcal{S}(\mathbb{R}_+^*)$ and such that

$$\int_{\hat{\mathbb{Z}}\times\mathbb{R}_+^*} f(\rho,v)\,d\rho\,dv=0\,. \tag{4.224}$$

Then the estimate

$$\mathbf{\Sigma}(\tilde f)(j):=\sum_{q\in\mathbb{Q}^*}\tilde f(qj)=O\left(|\log|j||^{-N}\right) \tag{4.225}$$

holds for $|j|\to 0$ and for all $N>0$.

PROOF. In (4.225) we are extending the function $f(\rho,v)=f(1,\rho,v)$ to adèles by zero as in (4.217). We can assume that there exists a character χ_n of the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^*$ such that

$$g(ma,v)=\chi_n(m)g(a,v),\qquad \forall a\in\mathbb{Z}/n\mathbb{Z},\quad\text{and}\quad\forall m\in(\mathbb{Z}/n\mathbb{Z})^*. \tag{4.226}$$

Let $\chi=\chi_n\circ p_n$. We then obtain

$$\tilde f(ma,v)=\chi(m)\tilde f(a,v),\qquad\forall a\in\hat{\mathbb{Z}},\ \ \forall m\in\hat{\mathbb{Z}}^*. \tag{4.227}$$

In particular this gives

$$\mathbf{\Sigma}(\tilde f)(mj)=\chi(m)\,\mathbf{\Sigma}(\tilde f)(j),\quad \forall m\in\hat{\mathbb{Z}}^*,\ \forall j\in C_{\mathbb{Q}}. \tag{4.228}$$

This shows that, in order to prove (4.225), it suffices to prove the estimate

$$|h(\lambda)|=O(|\log|\lambda|\,|^{-N}),\quad\text{for}\ \ \lambda\to 0, \tag{4.229}$$

where

$$h(\lambda)=\int_{\hat{\mathbb{Z}}^*}\mathbf{\Sigma}(\tilde f)(m\,j)\,\chi_0(m)\,d^*m, \tag{4.230}$$

for all j with $|j|=\lambda$, where $\chi_0=\bar\chi$ denotes the conjugate of χ. We write the Fourier transform of h in the form $\mathbf{F}_\mu$ over $\mathbb{R}_+^*$ as in (4.205), and we obtain

$$(\mathbf{F}_\mu h)(t)=\int_{\mathbb{R}_+^*}h(\lambda)\,\lambda^{-it}\,d^*\lambda=\int_{\mathbb{A}_{\mathbb{Q}}^*}\tilde f(j)\,\chi_0(j)\,|j|^s\,d^*j, \tag{4.231}$$

where $s=-it$ and d^*j is the multiplicative Haar measure on the idèles. The character χ_0 is extended as a unitary character of the idèles, which we still denote in the same way. More precisely, we identify $\mathbb{A}_{\mathbb{Q}}^*=\hat{\mathbb{Z}}^*\times\mathbb{R}_+^*\times\mathbb{Q}^*$ where $\mathbb{Q}^*$ is embedded as principal idèles. We extend χ_0 by 1 on $\mathbb{R}_+^*\times\mathbb{Q}^*\subset\mathbb{A}_{\mathbb{Q}}^*$. Thus χ_0 becomes a unitary Grössencharakter.

We can then apply the result of Weil on homogeneous distributions, Lemma 2.52 of Chapter 2 ([**296**], cf. [**71**] Appendix I, Lemma 1). For $\tilde f$ as above, this gives

$$\int_{\mathbb{A}_{\mathbb{Q}}^*}\tilde f(j)\,\chi_0(j)\,|j|^s\,d^*j=L(\chi_0,s)\,\langle\Delta'(s),\tilde f\rangle\,,\qquad\text{for}\ \ \Re(s)>1, \tag{4.232}$$

where $\Delta'=\otimes\Delta'_v$ is an infinite tensor product over the places of $\mathbb{Q}$ of homogeneous distributions on the local fields $\mathbb{Q}_v$ and $L(\chi_0,s)$ is the L-function with Grössencharakter χ_0. The identity (4.232) is a priori only valid for $\tilde f\in\mathcal{S}(\mathbb{A}_{\mathbb{Q}})$, but it remains valid for our $\tilde f$ by the absolute convergence of both sides.

In fact, we can assume that $f(\rho, v) = f_0(\rho) f_\infty(v)$, where $f_\infty \in \mathcal{S}(\mathbb{R}_+^*)$. One then gets $\tilde{f} = \tilde{f}_0 \otimes \tilde{f}_\infty$ and

$$\langle \Delta'(s), \tilde{f} \rangle = \langle \Delta'_0(s), \tilde{f}_0 \rangle \, \langle \Delta'_\infty(s), \tilde{f}_\infty \rangle, \tag{4.233}$$

where by construction (cf. (2.364) of Chapter 2) one has

$$\langle \Delta'_\infty(s), \tilde{f}_\infty \rangle = \int_{\mathbb{R}_+^*} f_\infty(\lambda) \, \lambda^s d^*\lambda \,, \quad d^*\lambda = d\lambda/\lambda. \tag{4.234}$$

Using (4.207), one gets $\tilde{\mathbf{F}}(f_\infty) \in \mathcal{S}(I_\beta)$. Thus, the function $t \mapsto \langle \Delta'_\infty(-it), \tilde{f}_\infty \rangle$ is in $\mathcal{S}(I_\beta)$. The term $\langle \Delta'_0(s), \tilde{f}_0 \rangle$ is simple to compute explicitly and one obtains that the function $t \mapsto \langle \Delta'_0(-it), \tilde{f}_0 \rangle$ is a holomorphic function on the strip I_β such that its restriction to the boundary ∂I_β is a tempered growth distribution, i.e. this function is in $\mathrm{Hol}(I_\beta)$, as in Definition 4.76.

Indeed, it is enough to do the computation when f_0 is a finite tensor product $f_0 = \otimes_{v \in S} f_p$ over a finite set S of primes, while f_p is a locally constant function on $\mathbb{Q}_p$ vanishing outside $\mathbb{Z}_p$ and such that $f_p(ma) = \chi(m) \, f(a)$, for all $a \in \mathbb{Z}_p$, $m \in \mathbb{Z}_p^*$, using (4.227) for the restriction of χ to $\mathbb{Z}_p^*$. One has

$$\langle \Delta'_0(s), \tilde{f}_0 \rangle = \prod_S \langle \Delta'_p(s), \tilde{f}_p \rangle.$$

One can assume that S contains the finite set P of places at which χ is ramified, i.e. where the restriction of χ to $\mathbb{Z}_p^*$ is non-trivial. For $p \in P$ one gets that f_p vanishes in a neighborhood of $0 \in \mathbb{Z}_p$, while

$$\langle \Delta'_p(s), \tilde{f}_p \rangle = \int_{\mathbb{Q}_p^*} f_p(x) \, \chi_0(x) \, |x|^s \, d^*x,$$

which, as a function of s, is a finite sum of exponentials p^{ks}. For $p \notin P$, one has the normalization condition $\langle \Delta'_p(s), 1_{\mathbb{Z}_p} \rangle = 1$ and the function f_p is a finite linear combination of the characteristic functions $x \mapsto 1_{\mathbb{Z}_p}(p^k x)$. It follows from homogeneity of the distribution $\Delta'_p(s)$ that $\langle \Delta'_p(s), \tilde{f}_p \rangle$, as a function of s, is a finite sum of exponentials of the form p^{ks}.

Similarly, if χ_0 is non-trivial, the function $L(\chi_0, -it)$ belongs to $\mathrm{Hol}(I_\beta)$ and $\zeta(-it) = L(1, -it)$ is the sum of $1/(-it-1)$ with an element of $\mathrm{Hol}(I_\beta)$. The presence of the pole at $t = i$ for $\chi_0 = 1$ is taken care of by the condition (4.224). All of this shows that

$$\mathbf{F}_\mu h \in \mathcal{S}(I_\beta), \quad \text{and } \mathbf{F}_\mu h(z) = 0 \quad \forall z \in I_\beta \quad \text{with } L(\chi_0, -iz) = 0. \tag{4.235}$$

This shows in particular that $h \in \mathcal{S}(\mathbb{R}_+^*)$, which gives the required estimate (4.229) on $|h(\lambda)|$, for $\lambda \to 0$. ∎

Remark 4.90. One could take instead of the extension $\tilde{f}$ a more symmetric choice where f is extended by 0 outside $\hat{\mathbb{Z}} \times \mathbb{R}^*$ and one requires the parity

$$\tilde{f}(-u, -\lambda) = f(u, \lambda) \,. \tag{4.236}$$

This choice introduces a factor of 2 in $\mathbf{\Sigma}(\tilde{f})$ and the above formulas remain valid provided one replaces (4.234) by

$$\langle \Delta'_\infty(s), \tilde{f}_\infty \rangle = \int_{\mathbb{R}^*} f_\infty(\lambda) \, \chi_\infty(\lambda) \, \lambda^s d^*\lambda \,, \tag{4.237}$$

where χ_∞ is the restriction of χ_0 to $\mathbb{R}^*$. This restriction is the character given by the sign of λ when the value of χ_0 on $-1 \in \hat{\mathbb{Z}}^*$ is -1, in which case the function $\tilde{f}_\infty$ is odd.

We can then prove the following result.

THEOREM 4.91. *Let $(\mathcal{A}, \sigma)$ be the Bost–Connes system. For all $\beta > 1$ and for $X = \int x(t)\, U_t\, dt \in \hat{\mathcal{A}}_\beta$, one obtains for $c \to -\infty$ the estimate*

$$\operatorname{Tr}(\pi_{\varepsilon, H+c}(X)) = e^{-c}\tau(X) + O(|c|^{-N}), \quad \forall N > 0, \tag{4.238}$$

where τ is the canonical trace on $\hat{\mathcal{A}}$ dual to the KMS$_1$ state φ on $(\mathcal{A}, \sigma)$.

PROOF. If φ denotes the KMS$_1$ state (4.126) for the Bost–Connes system, then the dual trace $\tau = \tau_\varphi$ is given by (4.222). Thus, we need to show that, for any $u \in \hat{\mathbb{Z}}^*$ and any function $f(k, \rho, \lambda)$ in $\iota\hat{\mathcal{A}}_\beta$, one has an estimate for $\lambda \to 0$ of the form

$$\sum_{n \in \mathbb{N}} f(1, nu, n\lambda) = \lambda^{-1} \int_{\hat{\mathbb{Z}} \times \mathbb{R}_+^*} f(1, \rho, v)\, d\rho\, dv + O(|\log(\lambda)|^{-N}), \tag{4.239}$$

for all $N > 0$. We use Theorem 4.85 above, by extending the function $f(\rho, v) = f(1, \rho, v)$ to adèles, and we write the left hand side of (4.239) as in (4.217).

If the extended function $\tilde{f}$ were in the Bruhat–Schwartz space $\mathcal{S}(\mathbb{A}_\mathbb{Q})$, then one could use the Poisson summation formula

$$|j| \sum_{q \in \mathbb{Q}} h(q\, j) = \sum_{q \in \mathbb{Q}} \hat{h}(q\, j^{-1}),$$

with $\hat{h}$ the Fourier transform of h, in order to get a better estimate with a remainder in $O(|\lambda|^N)$. However, in general one has $\tilde{f} \notin \mathcal{S}(\mathbb{A}_\mathbb{Q})$, because of the singularity at $\lambda = 0$. At least some functions f with non-zero integral do extend to an element $\tilde{f} \in \mathcal{S}(\mathbb{A}_\mathbb{Q})$ and this allows us to restrict to functions of two variables $f(\rho, v) = f(1, \rho, v)$ such that

$$\int_{\mathbb{R}_+^*} f(\rho, v)\, d\rho\, dv = 0.$$

We know by Lemma 4.83 that the function has finite level in $\rho \in \hat{\mathbb{Z}}$ and rapid decay in $v \in \mathbb{R}_+^*$. We can then apply the result of Proposition 4.89. That provides the estimate that completes the proof of the theorem. □

An immediate consequence of the estimate of Theorem 4.91 is the following result.

COROLLARY 4.92. *Let $(\mathcal{A}, \sigma)$ be the BC system. Let $X = \int x(t) U_t\, dt$ be an element of $\hat{\mathcal{A}}_\beta$ with $\tau(X) = 0$, i.e. in the kernel of the dual trace. Then the Fourier transform*

$$\mathbf{F}(f_\varepsilon)(u) = \int_\mathbb{R} f_\varepsilon(c)\, e^{-icu}\, dc \tag{4.240}$$

of the function $f_\varepsilon(c) = \operatorname{Tr}(\pi_{\varepsilon, H+c}(X))$ is holomorphic in the strip I_β, bounded and smooth on the boundary. In fact $\mathbf{F}(f_\varepsilon) \in \mathcal{S}(I_\beta)$.

PROOF. Combining Theorem 4.91 with Lemma 4.70, we obtain that, in the BC case and for $x \in \hat{\mathcal{A}}_\beta$, $\tau(x) = 0$, the function $f_\varepsilon(c) = \mathrm{Tr}(\pi_{\varepsilon,H+c}(X))$ satisfies, for all $N > 0$, the decay conditions

$$
\begin{aligned}
|f_\varepsilon(c)| &= O(e^{-\beta\, c}\, |c|^{-N}), \quad \text{for} \quad c \to +\infty \\
|f_\varepsilon(c)| &= O(|c|^{-N}) \qquad\quad\;\, \text{for} \quad c \to -\infty\,.
\end{aligned}
\tag{4.241}
$$

This is the correct decay at $c \to \pm\infty$ in order to have a Fourier transform (in the variable c) that is holomorphic in the strip I_β.
To show that this can be further improved to $\mathbf{F}(f_\varepsilon) \in \mathcal{S}(I_\beta)$, we use the general result of Proposition 4.77. This makes it possible to improve the estimate (4.241) for the BC system immediately. Indeed, the function $h(s) = s$ belongs to $\mathrm{Hol}(I_\beta)$ and this shows that (4.241) holds for all derivatives of f_ε. $\square$

Summarizing, the results on the decay estimates for the distillation of the BC endomotive combine to give the following result.

COROLLARY 4.93. *In the BC system $(\mathcal{A}, \sigma)$, for any $\beta > 1$ and any element $X \in \hat{\mathcal{A}}_{\beta,\tau} = \hat{\mathcal{A}}_\beta \cap \mathrm{Ker}(\tau)$ one has*

$$
\mathbf{F} \circ \mathrm{Tr} \circ \pi(X) \in C(\Omega_\beta, \mathcal{S}(I_\beta)). \tag{4.242}
$$

When expressed in the variable $\lambda = e^c$, the regularity of the function $\mathrm{Tr} \circ \pi(X)$ can be written equivalently in the form $\mathrm{Tr} \circ \pi(X) \in C(\Omega_\beta, \mathcal{S}_\beta(\mathbb{R}_+^*))$ where, with $\mu(\lambda) = \lambda$, $\forall \lambda \in \mathbb{R}_+^*$, we let

$$
\mathcal{S}_\beta(\mathbb{R}_+^*) = \cap_{[0,\beta]}\, \mu^{-\alpha}\, \mathcal{S}(\mathbb{R}_+^*) \tag{4.243}
$$

be the intersection of the shifted Schwartz spaces of $\mathbb{R}_+^* \sim \mathbb{R}$.

DEFINITION 4.94. *Let $\mathcal{S}^\natural(\tilde{\Omega}_\beta)$ be the cyclic submodule of $C(\tilde{\Omega}_\beta)^\natural$ whose elements are functions with restriction to the main diagonal that belongs to $C(\Omega_\beta, \mathcal{S}_\beta(\mathbb{R}_+^*))$ with $\mathcal{S}_\beta(\mathbb{R}_+^*)$ defined as in* (4.243).

We then obtain the following result.

PROPOSITION 4.95. *Let $(\mathcal{A}, \varphi)$ be the Bost–Connes endomotive, with $\tau = \tau_\varphi$ the dual trace on $\hat{\mathcal{A}} = \mathcal{A} \rtimes_{\sigma_t^\varphi} \mathbb{R}$. Let $\hat{\mathcal{A}}^\natural_{\beta,\tau} \subset \hat{\mathcal{A}}^\natural_\beta$ be the cyclic module defined by* $\mathrm{Ker}(\tau)$ *as in Definition 4.74. Then the map* $\mathrm{Tr} \circ \pi$ *defines a cyclic morphism*

$$
\delta : \hat{\mathcal{A}}^\natural_{\beta,\tau} \to \mathcal{S}^\natural(\tilde{\Omega}_\beta). \tag{4.244}
$$

DEFINITION 4.96. *We define the distilled Λ-module $D_\tau(\mathcal{A}, \varphi)$ as the cokernel of the cyclic morphism $\delta = (\mathrm{Tr} \circ \pi)^\natural$ of* (4.244)*, namely*

$$
D_\tau(\mathcal{A}, \varphi) := \mathrm{Coker}(\delta : \hat{\mathcal{A}}^\natural_{\beta,\tau} \to \mathcal{S}^\natural(\tilde{\Omega}_\beta)). \tag{4.245}
$$

4.10. Spectral realization.

In this section we show how one gets the spectral realization of the zeros of the Riemann zeta function and of L-functions with Grössencharakter from the distillation of the Bost–Connes endomotive.

The cyclic morphism δ of (4.244) is equivariant with respect to the $\mathbb{R}_+^*$ action, as in Lemma 4.81. Thus, we consider the induced $\mathbb{R}_+^*$ representation

$$\theta_\lambda \in \mathrm{Aut}(HC_0(D_\tau(\mathcal{A},\varphi))), \qquad \forall \lambda \in \mathbb{R}_+^*. \tag{4.246}$$

In fact, since we are dealing with the distilled system obtained from the Bost–Connes endomotive, we can refine (4.246) by taking into account the fact that endomotives are naturally endowed also with a Galois action, in this case of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ factoring through the abelianization, cf. §2.7 above.

PROPOSITION 4.97. *Let $(\mathcal{A},\varphi)$ be the BC endomotive. Let χ be a character of the compact group $\hat{\mathbb{Z}}^*$.*

(1) *The character χ determines an idempotent $e_\chi \in \mathrm{End}_{\mathcal{EV}^0_{\mathbb{Q},\mathbb{C}}}(\mathcal{A}_\mathbb{Q})$*

$$e_\chi = \int_{\hat{\mathbb{Z}}^*} \chi(g)\, Z_g\, dg \tag{4.247}$$

in the (pseudo)-abelian category $\mathcal{EV}^0_{\mathbb{Q},\mathbb{C}}$ of algebraic endomotives, where

$$Z_g = \{(xg^{-1}x)\}$$

is the graph of g as a correspondence.

(2) *The expression*

$$p_\chi = \int_{\hat{\mathbb{Z}}^*} g\,\chi(g)\, dg \tag{4.248}$$

determines an idempotent p_χ in $\mathrm{End}_\Lambda(D_\tau(\mathcal{A},\varphi))$ in the abelian category of Λ-modules.

PROOF. (1) Recall that, by Proposition 4.30, the algebra $\mathcal{A}_\mathbb{Q} = A \rtimes S$ of the algebraic endomotive of the BC system is globally invariant under the action of $G = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ and this action factors through $\mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q})$. In particular, this means that we can consider the action of $g \in \mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q})$ on $X = \varprojlim_\alpha X_\alpha$ and the graph Z_g of g as a correspondence in $X \times X$. The action of G commutes with the semigroup action of $S = \mathbb{N}$, hence Z_g defines a correspondence in $\mathrm{End}_{\mathcal{EV}^0_{\mathbb{Q},\mathbb{C}}}(\mathcal{A}_\mathbb{Q})$ and so does (4.247). One has $e_\chi^2 = \int_{\hat{\mathbb{Z}}^*\times\hat{\mathbb{Z}}^*} gg'\,\chi(gg')\,dg\,dg' = e_\chi$.

(2) The Galois group $G^{\mathrm{ab}} \cong \hat{\mathbb{Z}}^*$ acts on $(\mathcal{A},\sigma)$ compatibly with the time evolution. One obtains an action on $\hat{\mathcal{A}}$ preserving $\hat{\mathcal{A}}_\beta$ and the kernel $\mathrm{Ker}(\tau)$ since the state φ is G^{ab}-invariant. Similarly, one has an action of G^{ab} on KMS_β states Ω_β, which is compatible with the action as symmetries of $(\mathcal{A},\varphi)$ by the intertwining property (cf. Proposition 4.30 and Chapter 3 above). Thus, the induced morphism $\delta : \hat{\mathcal{A}}^\natural_{\beta,\tau} \to \mathcal{S}^\natural(\tilde{\Omega}_\beta)$ is also G^{ab}-equivariant and one obtains an induced linear action of G^{ab} on $D_\tau(\mathcal{A},\varphi)$. One can then form the combination p_χ as in (4.248) and one sees as above that $p_\chi^2 = p_\chi$ so that we have an idempotent in $\mathrm{End}_\Lambda(D_\tau(\mathcal{A},\varphi))$. $\square$

COROLLARY 4.98. *Let χ range over the characters of $\hat{\mathbb{Z}}^*$. There is a corresponding direct sum decomposition of the representation* (4.246) *as*

$$\oplus_\chi HC_0(p_\chi D_\tau(\mathcal{A},\varphi)). \tag{4.249}$$

PROOF. The result simply follows from Proposition 4.97 and the fact that the idempotents p_χ add up to the identity, namely

$$\sum_\chi p_\chi = \mathrm{Id}. \tag{4.250}$$

$\square$

We now show that the representation of $\mathbb{R}_+^*$ on

$$\mathcal{E} := HC_0(p_\chi\, D(\mathcal{A},\varphi)) \tag{4.251}$$

gives the spectral realization of the zeros of the L-function L_χ. More precisely, we have the following result.

THEOREM 4.99. *Let $z \in I_\beta$ be viewed as a character of* $\mathrm{Hol}(I_\beta)$ *and $\mathbb{C}_z$ be the corresponding one dimensional module. One has*

$$\mathcal{E} \otimes_{\mathrm{Hol}(I_\beta)} \mathbb{C}_z \neq \{0\} \Longleftrightarrow L_\chi(-iz) = 0. \tag{4.252}$$

PROOF. It follows from Proposition 4.89 and (4.235) that all elements h of the range of δ in degree zero have Fourier transform $\mathbf{F}_\mu h \in \mathcal{S}(I_\beta)$ that vanishes at any z such that $L(\chi_0, -iz) = 0$.
Moreover, explicit choices of the test function show that, if we let $L_{\mathbb{Q}}$ be the complete L-function (with its Archimedean Euler factors), then, for χ_0 non-trivial, $L_{\mathbb{Q}}(\chi_0, -iz)$ is the Fourier transform of an element of the range of δ.
When χ_0 is trivial, the range of δ contains $\xi(z - \frac{i}{2})$, where ξ is the Riemann ξ function (cf. Chapter 2 above and [**254**])

$$\xi(t) := -\frac{1}{2}\left(\frac{1}{4}+t^2\right)\Gamma\left(\frac{1}{4}+\frac{it}{2}\right)\pi^{-\frac{1}{4}-\frac{it}{2}}\zeta\left(\frac{1}{2}+it\right).$$

Thus, we see that, for h in the range of δ in degree zero, the set $\mathcal{Z}_{\chi_0}$ of common zeros of the $\mathbf{F}_\mu h$ is given by

$$\mathcal{Z}_{\chi_0} = \{z \in I_\beta \mid L_{\mathbb{Q}}(\chi_0, -iz) = 0\}. \tag{4.253}$$

Using Proposition 4.77, one then gets the required spectral realization. In fact, Proposition 4.77 shows that the $\mathrm{Hol}(I_\beta)$-module $\mathcal{E}$ of (4.251) is the quotient of $\mathcal{S}(I_\beta)$ by the $\mathrm{Hol}(I_\beta)$-submodule $\mathcal{N} = \{\mathbf{F}_\mu h \mid h \in \delta(\hat{\mathcal{A}}_\beta)\}$. Let $z \in I_\beta$ with $L_{\mathbb{Q}}(\chi_0, -iz) = 0$. Then the map $\mathcal{S}(I_\beta) \ni f \mapsto f(z)$ vanishes on $\mathcal{N}$ and yields a non-trivial map of $\mathcal{E} \otimes_{\mathrm{Hol}(I_\beta)} \mathbb{C}_z$ to $\mathbb{C}$. Let $z \in I_\beta$ with $L_{\mathbb{Q}}(\chi_0, -iz) \neq 0$. Any element of $\mathcal{E} \otimes_{\mathrm{Hol}(I_\beta)} \mathbb{C}_z$ is of the form $f \otimes 1$, with $f \in \mathcal{S}(I_\beta)$. There exists an element $\xi = g \otimes 1 \in \mathcal{N}$ such that $g(z) \neq 0$. Thus, we can assume, since we work modulo $\mathcal{N}$, that $f(z) = 0$. It remains to show that $f \otimes 1$ is in $\mathcal{S}(I_\beta) \otimes_{\mathrm{Hol}(I_\beta)} \mathbb{C}_z$. To see this, one writes $f(t) = (t - z)\, f_1(t)$ and one checks that $f_1 \in \mathcal{S}(I_\beta)$, while

$$f \otimes 1 = (t - z)\, f_1 \otimes 1 \sim f_1 \otimes (t - z) = 0.$$

$\square$

5. A cohomological Lefschetz trace formula

In the previous sections we developed the general formalism of the cooling map

$$\pi : \hat{\mathcal{A}}_\beta \to C(\tilde{\Omega}_\beta, \mathcal{L}^1)$$

and the corresponding cyclic morphism

$$(\mathrm{Tr} \circ \pi)^\natural : \hat{\mathcal{A}}_\beta^\natural \to C(\tilde{\Omega}_\beta)^\natural$$

with the distilled Λ-module $D(\mathcal{A}_{\mathbb{C}}, \varphi)$ given by the cokernel of this morphism.

In particular, we showed that, in the case of the Bost–Connes endomotive, the cyclic morphism

$$(\mathrm{Tr} \circ \pi)^\natural : \hat{\mathcal{A}}_{\beta,\tau}^\natural \to \mathcal{S}^\natural(\tilde{\Omega}_\beta)$$

gives the map $\mathfrak{E}$ of Chapter 2 §6 (cf. [**71**]) and gives a spectral realization of the zeros of the Riemann zeta function and L-functions with Grössencharakter from the scaling action of $\mathbb{R}_+^*$ on the cyclic homology $HC_0(p_\chi D_\tau(\mathcal{A}, \varphi))$.

We now show that the case of the Bost–Connes endomotive can be seen as a special case of a more general "restriction map" from the adèle class space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ of a global field to the idèle class group $C_{\mathbb{K}}$. The main result is a trace formula for the action of $C_{\mathbb{K}}$ on the HC_0 of the cokernel of the corresponding cyclic morphism, which recovers the Weil explicit formula and reformulates the RH as a positivity problem (in this context we are not working in the Hilbert space setting).

5.1. The adèle class space of a global field.

Let $\mathbb{K}$ be a global field, that is, either a number field (a finite algebraic extension of $\mathbb{Q}$) or a function field (the field $\mathbb{F}_q(C)$ of functions of a smooth projective curve over a finite field $\mathbb{F}_q$). Recall that we introduced in Chapter 2, Definition 2.45, the adèle class space of $\mathbb{K}$ as the quotient $X_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$. It is clear already by what we have seen in the case of $\mathbb{Q}$ that this quotient is better treated using the methods of noncommutative geometry, due to the ergodic nature of the action which makes the ordinary quotient ill behaved topologically. This means that we want to describe the ring of functions of the space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ using a noncommutative algebra. The natural choice of the algebra is the crossed product

$$C_0(\mathbb{A}_{\mathbb{K}}) \rtimes \mathbb{K}^* \quad \text{with the smooth subalgebra} \quad \mathcal{S}(\mathbb{A}_{\mathbb{K}}) \rtimes \mathbb{K}^*. \tag{4.254}$$

Elements of the smooth algebra (4.254) are finite sums

$$\sum_{k \in \mathbb{K}^*} f_k\, U_k, \qquad \text{with} \quad f_k \in \mathcal{S}(\mathbb{A}_{\mathbb{K}}), \tag{4.255}$$

with the product rule specified by

$$(U_k\, f\, U_k^*)(x) = f(k^{-1}x), \tag{4.256}$$

for $f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})$, $k \in \mathbb{K}^*$, and $x \in \mathbb{A}_{\mathbb{K}}$.

Consider the groupoid law on $\mathcal{G}_{\mathbb{K}} = \mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K}}$ given by

$$(k, x) \circ (k', y) = (kk', y), \quad \forall k, k' \in \mathbb{K}^*, \quad \text{and } \forall x, y \in \mathbb{A}_{\mathbb{K}} \ \text{ with } x = k'y. \tag{4.257}$$

The composition (4.257) is defined whenever the source $s(k, x) = x$ agrees with the range $r(k', y) = k'y$.

LEMMA 4.100. *The algebras* (4.254) *are, respectively, the groupoid C^*-algebra $C^*(\mathcal{G}_{\mathbb{K}})$ and its dense subalgebra $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$.*

PROOF. The product in the groupoid algebra is given by the associative convolution product

$$(f_1 * f_2)(k, x) = \sum_{s \in \mathbb{K}^*} f_1(k\, s^{-1}, s\, x) f_2(s, x), \tag{4.258}$$

and the adjoint is given by $f^*(k, x) = \overline{f(k^{-1}, k\, x)}$.
The functions (on the groupoid) associated to $f \in \mathcal{S}(\mathbb{A}_K)$ and U_k are given, respectively, by

$$\begin{aligned} \tilde{f}(1, x) = f(x) \quad &\text{and} \quad f(k, x) = 0 \quad \forall k \neq 1 \\ U_k(k, x) = 1 \quad &\text{and} \quad U_g(k, x) = 0 \quad \forall g \neq k. \end{aligned} \tag{4.259}$$

One checks that the convolution product (4.258) gives

$$(U_k * \tilde{f} * U_k^*)(1, x) = f(k^{-1}x),$$

while the algebra isomorphism is obtained by the map,

$$\sum_{k \in \mathbb{K}} f_k\, U_k \mapsto g \in \mathcal{S}(\mathcal{G}_{\mathbb{K}}), \quad g(k, x) = f_k(kx), \quad \forall (k, x) \in \mathcal{G}_{\mathbb{K}}. \tag{4.260}$$

□

5.2. The cyclic module of the adèle class space.

As in §1.2, we can associate to the algebra $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ of the adèle class space an object in the category of Λ-modules.
We consider the cyclic module $\mathcal{S}(\mathcal{G}_{\mathbb{K}})^\natural$ and the two cyclic morphisms

$$\varepsilon_j : \mathcal{S}(\mathcal{G}_{\mathbb{K}})^\natural \to \mathbb{C} \tag{4.261}$$

given by

$$\varepsilon_0(\sum f_k\, U_k) = f_1(0) \quad \text{and} \quad \varepsilon_1(\sum f_k\, U_k) = \int_{\mathbb{A}_{\mathbb{K}}} f_1(x)\, dx \tag{4.262}$$

and in higher degree by

$$\varepsilon_j^\natural(a^0 \otimes \cdots \otimes a^n) = \varepsilon_j(a^0 \cdots a^n). \tag{4.263}$$

The morphism ε_1 is given by integration on $\mathbb{A}_{\mathbb{K}}$ with respect to the additive Haar measure. This is $\mathbb{K}^*$ invariant hence it defines a trace on $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$. In the case of $\mathbb{K} = \mathbb{Q}$ that we discussed in the previous sections, we have seen in Lemma 4.88 that this corresponds to the dual trace τ_φ (cf. Lemma 4.73) for the KMS$_1$-state φ associated to the time evolution of the BC system. The morphism ε_0 here takes into account the fact that we are imposing a vanishing condition at $0 \in \mathbb{A}_{\mathbb{K}}$. In fact, the Λ-module we associate to $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ is given by

$$\mathcal{S}(\mathcal{G}_{\mathbb{K}})_0^\natural := \operatorname{Ker} \varepsilon_0^\natural \cap \operatorname{Ker} \varepsilon_1^\natural. \tag{4.264}$$

Since the algebra $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ is non-unital our treatment of cyclic modules here follows the setting described in §1.3 above (cf. also the analogous situation in §4.8).

REMARK 4.101. *One can recover the norm $|\cdot| : C_{\mathbb{K}} \to \mathbb{R}_+^*$ from the range of the cyclic morphism $\varepsilon_1^\natural$ through*

$$\int_{\mathbb{A}_{\mathbb{K}}} f(\gamma^{-1}x)\, dx = |\gamma| \int_{\mathbb{A}_{\mathbb{K}}} f(x)\, dx. \tag{4.265}$$

The range of $\varepsilon_0^\natural$ gives the trivial representation of $C_{\mathbb{K}}$.

5.3. The restriction map to the idèle class group.

Consider the idèles $\mathbb{A}_{\mathbb{K}}^* = \mathrm{GL}_1(\mathbb{A}_{\mathbb{K}})$ of $\mathbb{K}$ with their natural locally compact topology induced by the map

$$\mathbb{A}_{\mathbb{K}}^* \ni g \mapsto (g, g^{-1}). \tag{4.266}$$

We can see the idèle class group $C_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^*$ as a subspace of the adèle class space in the following way.

LEMMA 4.102. *The pairs $((k,x),(k',y)) \in \mathcal{G}_{\mathbb{K}}$ such that both x and y are in $\mathbb{A}_{\mathbb{K}}^*$ form a full subgroupoid of $\mathcal{G}_{\mathbb{K}}$ which is isomorphic to $\mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K}}^*$.*

PROOF. Elements of $\mathbb{A}_{\mathbb{K}}$ whose orbit under the $\mathbb{K}^*$ action contains an idèle are also idèles. Thus, we obtain a groupoid that is a full subcategory of $\mathcal{G}_{\mathbb{K}}$. ☐

This implies the existence of a restriction map. Consider the map

$$\rho : \mathcal{S}(\mathbb{A}_K) \ni f \mapsto f|_{\mathbb{A}_{\mathbb{K}}^*}. \tag{4.267}$$

We denote by $C_\rho(\mathbb{A}_{\mathbb{K}}^*) \subset C(\mathbb{A}_{\mathbb{K}}^*)$ the range of ρ.

COROLLARY 4.103. *The restriction map ρ of* (4.267) *extends to an algebra homomorphism*

$$\rho : \mathcal{S}(\mathcal{G}_{\mathbb{K}}) \to C_\rho(\mathbb{A}_{\mathbb{K}}^*) \rtimes \mathbb{K}^*. \tag{4.268}$$

PROOF. The map (4.267) induced by the inclusion $\mathbb{A}_{\mathbb{K}}^* \subset \mathbb{A}_{\mathbb{K}}$ is continuous and $\mathbb{K}^*$ equivariant, hence the map

$$\rho(\sum_{k\in\mathbb{K}^*} f_k\, U_k) = \sum_{k\in\mathbb{K}^*} \rho(f_k)\, U_k$$

is an algebra homomorphism. ☐

The action of $\mathbb{K}^*$ on $\mathbb{A}_{\mathbb{K}}^*$ is free and proper so that we have an equivalence of the locally compact groupoids $\mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K}}^*$ and $\mathbb{A}_{\mathbb{K}}^*/\mathbb{K}^* = C_{\mathbb{K}}$. We use the exact sequence of locally compact groups

$$1 \to \mathbb{K}^* \to \mathbb{A}_{\mathbb{K}}^* \xrightarrow{p} C_{\mathbb{K}} \to 1 \tag{4.269}$$

to parameterize the orbits of $\mathbb{K}^*$ as the fibers $p^{-1}(x)$ for $x \in C_{\mathbb{K}}$. By construction the Hilbert spaces

$$\mathcal{H}_x = \ell^2(p^{-1}(x))\,, \quad \forall x \in C_K \tag{4.270}$$

form a continuous field of Hilbert spaces over $C_{\mathbb{K}}$. We let $\mathcal{L}^1(\mathcal{H}_x)$ be the Banach algebra of trace class operators in $\mathcal{H}_x$. These form a continuous field over $C_{\mathbb{K}}$.

PROPOSITION 4.104. *The restriction map ρ of* (4.267) *extends to an algebra homomorphism*

$$\rho : \mathcal{S}(\mathcal{G}_{\mathbb{K}}) \to C(C_{\mathbb{K}}, \mathcal{L}^1(\mathcal{H}_x))\,. \tag{4.271}$$

PROOF. Each $p^{-1}(x)$ is globally invariant under the action of $\mathbb{K}^*$ so the crossed product rules in $C_\rho(\mathbb{A}_{\mathbb{K}}^*) \rtimes \mathbb{K}^*$ are just multiplication of operators in $\mathcal{H}_x$. To show that the obtained operators are in $\mathcal{L}^1$ we just need to consider monomials $f_k\, U_k$. In that case the only non-zero matrix elements correspond to $k = xy^{-1}$. It is enough to show that, for any $f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})$, the function $k \mapsto f(k\, b)$ is summable. This follows from the discreteness of $b\,\mathbb{K} \subset \mathbb{A}_{\mathbb{K}}$ and the construction of the Bruhat–Schwartz space $\mathcal{S}(\mathbb{A}_{\mathbb{K}})$, cf. [**71**]. In fact the associated operator is of finite rank when f has compact support. In the general case the summability of the function $k \mapsto f(k\, b)$ follows as in the proof of Lemma 2.51 of Chapter 2. □

In general the exact sequence (4.269) does not split and one does not have a natural $C_{\mathbb{K}}$-equivariant trivialization of the continuous field $\mathcal{H}_x$. Thus it is important in the general case to keep the nuance between the algebras $C(C_{\mathbb{K}}, \mathcal{L}^1(\mathcal{H}_x))$ and $C(C_{\mathbb{K}})$. We shall first deal with the special case $\mathbb{K} = \mathbb{Q}$ in which this issue does not arise.

5.4. The Morita equivalence and cokernel for $\mathbb{K} = \mathbb{Q}$.

The exact sequence (4.269) splits for $\mathbb{K} = \mathbb{Q}$ and admits a natural continuous section which corresponds to the open and closed fundamental domain $\Delta_{\mathbb{Q}} = \hat{\mathbb{Z}}^* \times \mathbb{R}_+^* \subset \mathbb{A}_{\mathbb{Q}}^*$ for the action of $\mathbb{Q}^*$ on idèles. This allows one to construct a cyclic morphism between the cyclic module associated to the algebra $C_\rho(\mathbb{A}_{\mathbb{Q}}^*) \rtimes \mathbb{Q}^*$ and a suitable algebra $C_\rho(C_{\mathbb{Q}})$ of functions on $C_{\mathbb{Q}}$.

LEMMA 4.105. *The composition $d_{\mathbb{Q}} \circ e_{\mathbb{Q}}$ of the maps*

$$e_{\mathbb{Q}} : (k, hb) \mapsto (b, (k, h)) \quad \text{and} \quad d_{\mathbb{Q}}(k, h) = (kh, h), \tag{4.272}$$

with $b \in \Delta_{\mathbb{Q}}$ and $k, h \in \mathbb{Q}^*$, gives an isomorphism of the locally compact groupoids

$$\mathbb{Q}^* \ltimes \mathbb{A}_{\mathbb{Q}}^* \simeq \Delta_{\mathbb{Q}} \times \mathbb{Q}^* \times \mathbb{Q}^*. \tag{4.273}$$

PROOF. The map $e_{\mathbb{Q}}$ realizes an isomorphism between the locally compact groupoids

$$\mathbb{Q}^* \ltimes \mathbb{A}_{\mathbb{Q}}^* \simeq \Delta_{\mathbb{Q}} \times (\mathbb{Q}^* \ltimes \mathbb{Q}^*),$$

where $\mathbb{Q}^* \ltimes \mathbb{Q}^*$ is the groupoid of the action of $\mathbb{Q}^*$ on itself by multiplication. The latter is isomorphic to the trivial groupoid $\mathbb{Q}^* \times \mathbb{Q}^*$ via the map $d_{\mathbb{Q}}$. □

We then have the following result.

PROPOSITION 4.106. *The map*

$$\sum_{k \in \mathbb{Q}^*} f_k\, U_k \mapsto M_b(x, y) = f_{xy^{-1}}(x\, b), \tag{4.274}$$

for $x, y \in \mathbb{Q}^*$ with $k = xy^{-1}$ and $b \in \Delta_{\mathbb{Q}}$, defines an algebra homomorphism

$$C_\rho(\mathbb{A}_{\mathbb{Q}}^*) \rtimes \mathbb{Q}^* \to C(\Delta_{\mathbb{Q}}, M_\infty(\mathbb{C}))$$

to the algebra of matrix-valued functions on $\Delta_{\mathbb{Q}}$. For any $f \in \mathcal{S}(\mathcal{G}_{\mathbb{Q}})$ the element M_b obtained in this way is of trace class.

PROOF. We use the groupoid isomorphism (4.272) to write $k = xy^{-1}$ and $khb = xb$, for $x = kh$ and $y = h$. The second statement follows from Proposition 4.104. □

DEFINITION 4.107. *In the number field case, the "strong Schwartz space" is defined as*

$$\mathbf{S}\,(C_{\mathbb{K}}) = \cap_{\beta\in\mathbb{R}}\, \mu^{\beta}\mathcal{S}(C_{\mathbb{K}}), \tag{4.275}$$

where $\mu \in C(C_{\mathbb{K}})$ is the module morphism from $C_{\mathbb{K}}$ to $\mathbb{R}_+^*$. In the function field case we let $\mathbf{S}\,(C_{\mathbb{K}})$ be the space of Schwartz functions with compact support.

Let π be the composition of the restriction map ρ of (4.268) with the algebra morphism (4.274). Since the trace Tr on $M_\infty(\mathbb{C})$ gives a cyclic morphism one can use this to obtain a morphism of cyclic modules $(\mathrm{Tr}\circ\pi)^\natural$, which we now describe explicitly.

PROPOSITION 4.108. *The map $\mathrm{Tr}\circ\pi$ defines a morphism $(\mathrm{Tr}\circ\pi)^\natural$ of cyclic modules from $\mathcal{S}(\mathcal{G}_{\mathbb{Q}})_0^\natural$ to the cyclic submodule $\mathbf{S}^{\,\natural}(C_{\mathbb{Q}}) \subset C(C_{\mathbb{Q}})^\natural$ whose elements are continuous functions whose restriction to the main diagonal belongs to $\mathbf{S}\,(C_{\mathbb{Q}})$.*

PROOF. By Proposition 4.106 the map π is an algebra homomorphism from $\mathcal{S}(\mathcal{G}_{\mathbb{Q}})$ to $C(\Delta_{\mathbb{Q}},\mathcal{L}^1) \sim C(C_{\mathbb{Q}},\mathcal{L}^1)$. We need to show that the corresponding cyclic morphism using $\mathrm{Tr}^\natural$ lands in the cyclic submodule $\mathbf{S}^{\,\natural}(C_{\mathbb{Q}})$.
For simplicity we can just restrict to the case of monomials, where we consider elements of the form

$$Z = f_{k_0}\,U_{k_0}\otimes f_{k_1}\,U_{k_1}\otimes\cdots\otimes f_{k_n}\,U_{k_n}. \tag{4.276}$$

The matrix valued functions associated to the monomials $f_{k_j}\,U_{k_j}$ as in Proposition 4.106 have matrix elements at a point $b\in\Delta_{\mathbb{Q}}$ that are nonzero only for $x_{j+1} = x_j k_j^{-1}$ and are of the form

$$f_{k_j}\,U_{k_j} \mapsto M_b(x_j, x_{j+1}) = f_{k_j}(x_j b). \tag{4.277}$$

Composing with the cyclic morphism $\mathrm{Tr}^\natural$ gives

$$(\mathrm{Tr}\circ\pi)^\natural(Z)(b_0, b_1, \cdots, b_n) = \sum\prod M_{b_j}(x_j, x_{j+1}) \tag{4.278}$$

where the $x_j\in\mathbb{K}^*$ and $x_{n+1} = x_0$. Let $\gamma_0 = 1$ and $\gamma_{j+1} = k_j\gamma_j$. Then we find that $(\mathrm{Tr}\circ\pi)^\natural(Z) = 0$, unless $\prod_j k_j = 1$, i.e. $\gamma_{n+1} = 1$. In this case we obtain

$$\mathrm{Tr}\circ\pi(Z)(b_0, b_1, \ldots, b_n) = \sum_{k\in\mathbb{Q}^*}\prod_{j=0}^{n} f_{k_j}(\gamma_j^{-1}kb_j), \quad \forall b_j\in\Delta_{\mathbb{Q}}. \tag{4.279}$$

When restricted to the main diagonal $b_j = b$ this gives

$$\mathrm{Tr}\circ\pi(Z)(b, b, \cdots, b) = \sum_{k\in\mathbb{Q}^*}\prod_{j=0}^{n} f_{k_j}(\gamma_j^{-1}kb), \quad \forall b\in\Delta_{\mathbb{Q}}\,, \tag{4.280}$$

For $n = 0$ the formula (4.279) reduces to the form

$$\mathbf{\Sigma}(f)(b) = \sum_{k\in\mathbb{Q}^*} f(kb), \quad \forall b\in\Delta_{\mathbb{Q}}, \quad f = \prod_{j=0}^{n} f_{k_j}\,U_{k_j}\,. \tag{4.281}$$

One has $f\in\mathcal{S}(\mathbb{A}_{\mathbb{Q}})_0$ (where $\mathcal{S}(\mathbb{A}_{\mathbb{Q}})_0 = \mathrm{Ker}\,\varepsilon_0\cap\mathrm{Ker}\,\varepsilon_1\subset\mathcal{S}(\mathbb{A}_{\mathbb{Q}})$) since $\varepsilon_j(Z) = 0$. Then $\mathbf{\Sigma}(f)$ is an element of $\mathbf{S}\,(C_{\mathbb{Q}})$, by Lemma 2.51 of Chapter 2. In general, (4.279) gives a continuous function of $n+1$ variables on $C_{\mathbb{Q}}$, and its restriction to the main diagonal belongs to $\mathbf{S}\,(C_{\mathbb{Q}})$. ☐

Since the category of cyclic modules is an abelian category, we can consider the cokernel in the category of Λ-modules of the cyclic morphism $(\mathrm{Tr}\circ\pi)^\natural$, with π the composite of (4.268) and (4.274). This works nicely for $\mathbb{K}=\mathbb{Q}$ but makes use of the splitting of the exact sequence (4.269).

5.5. The cokernel of ρ for general global fields.

To handle the general case in a canonical manner one just needs to work directly with $C(C_{\mathbb{K}},\mathcal{L}^1(\mathcal{H}_x))$ instead of $C(C_{\mathbb{K}})$ and express at that level the decay condition of the restrictions to the diagonal in the cyclic submodule $\mathbf{S}^\natural(C_{\mathbb{Q}})$ of Proposition 4.108.

DEFINITION 4.109. *We define $\mathbf{S}^\natural(C_{\mathbb{K}},\mathcal{L}^1(\mathcal{H}_x))$ to be the cyclic submodule of the cyclic module $C(C_{\mathbb{K}},\mathcal{L}^1(\mathcal{H}_x))^\natural$, whose elements are continuous functions such that the trace of the restriction to the main diagonal belongs to $\mathbf{S}(C_{\mathbb{K}})$.*

Note that for $T\in C(C_{\mathbb{K}},\mathcal{L}^1(\mathcal{H}_x))^\natural$ of degree n, $T(x_0,\dots,x_n)$ is an operator in $\mathcal{H}_{x_0}\otimes\cdots\otimes\mathcal{H}_{x_n}$. On the diagonal, $x_j=x$ for all j, the trace map corresponding to $\mathrm{Tr}^\natural$ is given by

$$\mathrm{Tr}^\natural(T_0\otimes T_1\otimes\cdots\otimes T_n)=\mathrm{Tr}(T_0\,T_1\,\cdots\,T_n)\,. \tag{4.282}$$

This makes sense since on the diagonal all the Hilbert spaces $\mathcal{H}_{x_j}$ are the same.

The argument of Proposition 4.108 extends to the general case and shows that the cyclic morphism $\rho^\natural$ of the restriction map ρ lands in $\mathbf{S}^\natural(C_{\mathbb{K}},\mathcal{L}^1(\mathcal{H}_x))$.

DEFINITION 4.110. *We define $\mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*,C_{\mathbb{K}})$ to be the cokernel of the cyclic morphism*

$$\rho^\natural\ :\ \mathcal{S}(\mathcal{G}_{\mathbb{K}})_0^\natural\to\mathbf{S}^\natural(C_{\mathbb{K}},\mathcal{L}^1(\mathcal{H}_x))$$

Moreover, an important issue arises, since the ranges of continuous linear maps are not necessarily closed subspaces. In order to preserve the duality between cyclic homology and cyclic cohomology we shall define the cokernel of a cyclic map $T:\mathcal{A}^\natural\to\mathcal{B}^\natural$ as the quotient of $\mathcal{B}^\natural$ by the closure of the range of T. In a dual manner, the kernel of the transposed map $T^t:\mathcal{B}^\sharp\to\mathcal{A}^\sharp$ is automatically closed and is the dual of the above.

The choice of the notation $\mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*,C_{\mathbb{K}})$ is explained by the fact that we consider this a first cohomology group, in the sense that it is a cokernel in a sequence of cyclic homology groups for the inclusion of the idèle class group in the adèle class space (dually for the restriction map of algebras), hence we can think of it as giving rise to an H^1 in the relative cohomology sequence of an inclusion of $C_{\mathbb{K}}$ in the noncommutative space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$.

When we discussed the cooling and distillation in §4.8 we introduced the cokernel $D(\mathcal{A}_{\mathbb{C}},\varphi)$ of the cooling map in the abelian category of cyclic modules and its cyclic homology $HC_0(D(\mathcal{A}_{\mathbb{C}},\varphi))$. We now proceed in a similar manner, using the description of [**65**] of cyclic (co)homology in terms of derived functors in the category of cylic modules, as in §1.4 above, and write the cyclic homology as

$$HC_n(\mathcal{A})=\mathrm{Tor}_n(\mathbb{C}^\natural,\mathcal{A}^\natural). \tag{4.283}$$

Thus, we obtain a cohomological realization of the cyclic module

$$\mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*,C_{\mathbb{K}})$$

by setting

$$H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}}) := \mathrm{Tor}(\mathbb{C}^\natural, \mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})). \tag{4.284}$$

As we observed before, we think of this as an H^1 because of its role as a relative term in a cohomology exact sequence of the pair $(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$.
We now show that $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ carries an action of $C_{\mathbb{K}}$, which we can view as the abelianization $W_{\mathbb{K}}^{\mathrm{ab}} \sim C_{\mathbb{K}}$ of the Weil group. This action is induced by the multiplicative action of $C_{\mathbb{K}}$ on $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ and on itself. This generalizes to global fields the action of $C_{\mathbb{Q}} = \hat{\mathbb{Z}}^* \times \mathbb{R}_+^*$ on $HC_0(D(\mathcal{A}, \varphi))$ for the Bost–Connes endomotive.

Proposition 4.111. *The cyclic modules $\mathcal{S}(\mathcal{G}_{\mathbb{K}})_0^\natural$ and $\mathbf{S}^\natural(C_{\mathbb{K}}, \mathcal{L}^1(\mathcal{H}_x))$ are endowed with an action of $\mathbb{A}_{\mathbb{K}}^*$ and the morphism $\rho^\natural$ is $\mathbb{A}_{\mathbb{K}}^*$-equivariant. This induces an action of $C_{\mathbb{K}}$ on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$.*

Proof. For $\gamma \in \mathbb{A}_{\mathbb{K}}^*$ one defines an action by automorphisms of the algebra $\mathcal{A} = \mathcal{S}(\mathcal{G}_{\mathbb{K}})$ by setting

$$\vartheta_{\mathrm{a}}(\gamma)(f)(x) := f(\gamma^{-1}x), \quad \text{for } f \in \mathcal{S}(\mathbb{A}_{\mathbb{K}}), \tag{4.285}$$

$$\vartheta_{\mathrm{a}}(\gamma)(\sum_{k\in\mathbb{K}^*} f_k\, U_k) := \sum_{k\in\mathbb{K}^*} \vartheta_{\mathrm{a}}(\gamma)(f_k)\, U_k\,. \tag{4.286}$$

This action is inner for $\gamma \in \mathbb{K}^*$ and induces an outer action

$$C_{\mathbb{K}} \to \mathrm{Out}(\mathcal{S}(\mathcal{G}_{\mathbb{K}}))\,. \tag{4.287}$$

Similarly, the continuous field $\mathcal{H}_x = \ell^2(p^{-1}(x))$ over $C_{\mathbb{K}}$ is $\mathbb{A}_{\mathbb{K}}^*$-equivariant for the action of $\mathbb{A}_{\mathbb{K}}^*$ on $C_{\mathbb{K}}$ by translations, and the equality

$$(V(\gamma)\xi)(y) := \xi(\gamma^{-1}\, y)\,, \quad \forall y \in p^{-1}(\gamma x)\,,\ \xi \in \ell^2(p^{-1}(x))\,, \tag{4.288}$$

defines an isomorphism $\mathcal{H}_x \xrightarrow{V(\gamma)} \mathcal{H}_{\gamma x}$. One obtains then an action of $\mathbb{A}_{\mathbb{K}}^*$ on $C(C_{\mathbb{K}}, \mathcal{L}^1(\mathcal{H}_x))$ by setting

$$\vartheta_{\mathrm{m}}(\gamma)(f)(x) := V(\gamma)\, f(\gamma^{-1}\, x)\, V(\gamma^{-1}), \quad \forall f \in C(C_{\mathbb{K}}, \mathcal{L}^1(\mathcal{H}_x))\,. \tag{4.289}$$

The morphism ρ is $\mathbb{A}_{\mathbb{K}}^*$-equivariant, so that one obtains an induced action on the cokernel $\mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$. This action is inner for $\gamma \in \mathbb{K}^*$ and thus induces an action of $C_{\mathbb{K}}$ on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$. ☐

We denote by

$$C_{\mathbb{K}} \ni \gamma \mapsto \underline{\vartheta_{\mathrm{m}}}(\gamma) \tag{4.290}$$

the induced action on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$.
One obtains a corresponding decomposition of $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ according to projectors associated to characters of $C_{\mathbb{K},1}$, as we saw in Proposition 4.97 for the BC endomotive. We obtain in this way the following decomposition.

Proposition 4.112. *Characters $\chi \in \widehat{C_{\mathbb{K},1}}$ determine a canonical direct sum decomposition*

$$\begin{aligned} H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}}) &= \textstyle\bigoplus_{\chi\in\widehat{C_{\mathbb{K},1}}} H^1_\chi(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}}) \\ H^1_\chi(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}}) &= \{\xi |\, \underline{\vartheta_{\mathrm{m}}}(\gamma)\, \xi = \chi(\gamma)\, \xi,\ \forall \gamma \in C_{\mathbb{K},1}\}. \end{aligned} \tag{4.291}$$

where $\underline{\vartheta_{\mathrm{m}}}(\gamma)$ denotes the induced action (4.290) *on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$.*

PROOF. The action of $\mathbb{A}_{\mathbb{K}}^*$ on $\mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ induces a corresponding action of $C_{\mathbb{K}}$ on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$. □

The same argument we described in §4.10 above for the distillation of the BC endomotive applies in this context and gives the following result.

PROPOSITION 4.113. *The induced representation of $C_{\mathbb{K}}$ on $H^1_\chi(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ gives the spectral realization of the zeros of the L-function with χ as Grössencharakter.*

This result is a variant of Corollary 2 of [**71**]. There is a crucial difference from [**71**] in that all zeros (including those not located on the critical line) now appear due to the choice of the function spaces. To see what happens it is simpler to deal with the dual spaces, i.e. to compute the cyclic cohomology HC^0. Its elements are cyclic morphisms T from $\mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ to $\mathbb{C}^\natural$ and they are determined by the map T^0 in degree 0. The cyclic morphism property then shows that T^0 defines a trace on $\mathbf{S}^\natural(C_{\mathbb{K}}, \mathcal{L}^1(\mathcal{H}_x))$ which vanishes on the range of $\rho^\natural$. The freeness of the action of $\mathbb{K}^*$ on $\mathbb{A}_{\mathbb{K}}^*$ then ensures that these traces are given by continuous linear forms on $\mathbf{S}(C_{\mathbb{K}})$ which vanish on the following subspace of $\mathbf{S}(C_{\mathbb{K}})$ which is the range of the restriction map, defined as follows.

DEFINITION 4.114. *Let $\mathcal{V} \subset \mathbf{S}(C_{\mathbb{K}})$ denote the range of the map* $\mathrm{Tr} \circ \rho$*, i.e.*

$$\mathcal{V} = \{h \in \mathbf{S}(C_{\mathbb{K}}) |\, h(x) = \sum_{k \in \mathbb{K}^*} \xi(kx), \;\; \textit{with } \xi \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0\}, \tag{4.292}$$

where $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0 = \mathrm{Ker}\,\varepsilon_0 \cap \mathrm{Ker}\,\varepsilon_1 \subset \mathcal{S}(\mathbb{A}_{\mathbb{K}})$.

We have seen above in the case $\mathbb{K} = \mathbb{Q}$ that the range of $\mathrm{Tr} \circ \rho$ is indeed contained in $\mathbf{S}(C_{\mathbb{K}})$. The general case follows from Lemma 2.51 of Chapter 2. Moreover, we have the following results about the action $\underline{\vartheta_{\mathrm{m}}}(\gamma)$, for $\gamma \in C_{\mathbb{K}}$, on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$. Suppose given $f \in \mathbf{S}(C_{\mathbb{K}})$. We define a corresponding operator

$$\underline{\vartheta_{\mathrm{m}}}(f) = \int_{C_{\mathbb{K}}} f(\gamma)\, \underline{\vartheta_{\mathrm{m}}}(\gamma)\, d^*\gamma, \tag{4.293}$$

acting on the complex vector space $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$. Here $d^*\gamma$ is the multiplicative Haar measure on $C_{\mathbb{K}}$. We have the following description of the action of $\underline{\vartheta_{\mathrm{m}}}(f)$.

LEMMA 4.115. *For $f \in \mathbf{S}(C_{\mathbb{K}})$, the action of the operator $\underline{\vartheta_{\mathrm{m}}}(f)$ of (4.293) on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ is the action induced on the quotient of $\mathbf{S}(C_{\mathbb{K}})$ by $\mathcal{V} \subset \mathbf{S}(C_{\mathbb{K}})$ of the action of $\vartheta_{\mathrm{m}}(f)$ on $\mathbf{S}(C_{\mathbb{K}})$ by convolution product*

$$\vartheta_{\mathrm{m}}(f)\xi(u) = \int \xi(g^{-1}u) f(g)\, d^*g = (f \star \xi)(u). \tag{4.294}$$

PROOF. One first shows that one can lift f to a function $\tilde f$ on $\mathbb{A}_{\mathbb{K}}^*$ such that

$$\sum_{k \in \mathbb{K}^*} \tilde f(kx) = f(x)$$

and that convolution by $\tilde f$, i.e.

$$\int \tilde f(\gamma) \vartheta_{\mathrm{a}}(\gamma) d^*\gamma$$

leaves $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ globally invariant. This means showing that that $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$ is stable under convolution by the lift of $\mathbf{S}(C_{\mathbb{K}})$. Then (4.294) follows directly from the definition of the actions (4.289), (4.286), (4.290) and the operator (4.293). □

For $f \in \mathbf{S}(C_{\mathbb{K}})$ and $\tilde{\chi}$ the extension of a character $\chi \in \widehat{C_{\mathbb{K},1}}$ to $C_{\mathbb{K}}$ as in Chapter 2 §8.3, we consider the Fourier transform (Chapter 2, (2.298))

$$\hat{f}(\tilde{\chi},\rho) = \int_{C_{\mathbb{K}}} f(u)\tilde{\chi}(u)\,|u|^{\rho}\,d^*u. \tag{4.295}$$

We let $N^{\perp} := \{z \in \mathbb{C} |\, \lambda^z = 1,\ \forall \lambda \in N\}$ as in Chapter 2, Definition 2.38, (2.300).
The operators $\underline{\vartheta}_{\rm m}(f)$ of (4.293) have the following properties, [**75**], which give the *spectral side* of the trace formula.

THEOREM 4.116. *For any $f \in \mathbf{S}(C_{\mathbb{K}})$, the operator $\underline{\vartheta}_{\rm m}(f)$ defined in* (4.293) *acting on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ is of trace class. The trace is given by*

$$\operatorname{Tr}(\underline{\vartheta}_{\rm m}(f)|H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})) = \sum_{\rho \in \mathbb{C}/N^{\perp} |\, L(\tilde{\chi},\rho)=0} \hat{f}(\tilde{\chi},\rho), \tag{4.296}$$

with $\hat{f}(\tilde{\chi},\rho)$ the Fourier transform (4.295).

In the function field case the decomposition (4.291) of Proposition 4.112 splits H^1 as a direct sum of finite dimensional spaces and there is no technical difficulty in proving Theorem 4.116. Thus we shall concentrate on the number field case, and our goal will be to exhibit a tractable decomposition of H^1 as a direct sum of finite dimensional spaces. This will be done by constructing, for each character $\chi \in \widehat{C_{\mathbb{K},1}}$, a tempered sequence of idempotents $\Pi(N)$ for $N \in \mathbb{Z}$ which split $H^1_{\chi}(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ as a direct sum of finite dimensional spaces.

For definiteness we handle the case $\mathbb{K} = \mathbb{Q}$, with the trivial character $\chi \in \widehat{C_{\mathbb{Q},1}}$, but the arguments apply in general.
We let I be a vertical strip of bounded width containing $\{z |\, \Re(z) \in [-1, 2]\}$. We need to remove the pole and the trivial zeros of the zeta function without altering the polynomial growth of both $|\zeta(z)|$ and $1/|\zeta(z)|$ on the boundary ∂I. This is simply obtained by defining

$$\zeta_I(z) = (z-1) \prod_{n \in \mathbb{N}, -2n \in I} (z+2n)^{-1}\,\zeta(z)\,. \tag{4.297}$$

We let Z_{ζ} be the set (with multiplicity) of non-trivial zeros of $\zeta(z)$.

LEMMA 4.117. *The following conditions are equivalent for $f \in \mathcal{S}(I)$,*
(1) *For any $z_0 \in Z_{\zeta}$ of order n one has $f(z) = O(z - z_0)^n$.*
(2) *There exists $h \in \mathcal{S}(I)$ with $f = \zeta_I\, h$.*
(3) *The function f belongs to the closure of the range of $\mathbf{\Sigma}$ in $\mathcal{S}(I)$.*

PROOF. (1) $\Rightarrow$ (2). The function $h = f/\zeta_I$ is holomorphic and is of rapid decay on the boundary ∂I of the strip since the function $1/\zeta_I$ is of tempered growth on ∂I.
(2) $\Rightarrow$ (3). By Lemma 2.50 of Chapter 2, the range of $\mathbf{\Sigma}$ contains all functions of the form

$$\zeta(z)\,\Delta_{\infty}(z)(k) = \zeta(z) \int_{\mathbb{R}} k(x)|x|^z\,d^*x\,, \tag{4.298}$$

where $k \in \mathcal{S}(\mathbb{R})$ is an even function with $k(0) = \int k(x)dx = 0$. By Proposition 2.24 of Chapter 2 §6.1 we can find, for any complex numbers $a_j \in \mathbb{C}$ an element g of the range of $\mathbf{\Sigma}$ such that $g(-2j) = a_j$ for all $j \in \mathbb{N}$ with $-2j \in I$. Using this we can assume that the given function $f \in \mathcal{S}(I)$ vanishes at all these points and hence is of the form $f = \zeta\, h$ with $h(1) = 0$. Thus it is enough to approximate h by functions of the form

$$\int_{\mathbb{R}} k(x)|x|^z\, d^*x$$

where $k \in \mathcal{S}(\mathbb{R})$ is an even function with $\int k(x)dx = 0$ and $k(x) = O(x^{n+1})$ for $x \to 0$ for all $n \in \mathbb{N}$ with $-n \in I$.
(3) $\Rightarrow$ (1). This follows from Lemma 2.52 of Chapter 2, and the continuity of the map from $\mathcal{S}(I)$ to the jet of a function at a point of Z_ζ. $\square$

We let $J \subset \mathcal{S}(I)$ be defined by the equivalent conditions of Lemma 4.117. It is a closed ideal by construction.
Given a subset Z of I with multiplicity and a holomorphic function f on I we denote for each $\rho \in Z$ of multiplicity $n_\rho > 0$ by $f^{(n_\rho)}(\rho)$ the jet of f up to $O(z-\rho)^{n_\rho}$. The evaluation map on Z is defined by

$$\mathcal{S}(I) \ni f \mapsto f|_Z = (f^{(n_\rho)}(\rho))_{\rho \in Z}\,. \tag{4.299}$$

We let $\mathbb{C}^{(n_\rho)} \sim \mathbb{C}^{n_\rho}$ be the space of dimension n_ρ of jets of order n_ρ.

Corollary 4.118. *The evaluation on Z_ζ defines an injective map from the cokernel H^1 to the space of sequences*

$$\rho \in Z_\zeta \mapsto g(\rho) \in \mathbb{C}^{(n_\rho)}$$

Proof. The restriction map is well defined and, by Lemma 4.117, its kernel coincides with the closure of the range of $\mathbf{\Sigma}$ in $\mathcal{S}(I)$. $\square$

Definition 4.119. *For any $\lambda \in \mathbb{C} \setminus \{1\}$, $\zeta_I(\lambda) \neq 0$ we let*

$$R_\lambda(z) = \frac{1}{\lambda - z}\left(1 - \frac{\zeta_I(z)}{\zeta_I(\lambda)}\right) \tag{4.300}$$

This function is regular at $z = \lambda$ and will play the role of the resolvent.

Lemma 4.120. *Let $\lambda \in \mathbb{C}$, $\zeta_I(\lambda) \neq 0$.*

(1) *The multiplication by R_λ agrees with the resolvent $\frac{1}{\lambda - z}$ modulo the ideal generated by ζ_I.*
(2) *The multiplication by R_λ is a multiplier of $\mathcal{S}(I)$ preserving the ideal $J \subset \mathcal{S}(I)$.*
(3) *Its size is controlled by $\frac{1}{|\zeta_I(\lambda)|}$.*

Proof. (1) One has

$$R_\lambda(z)\,(\lambda - z) = 1 - \frac{\zeta_I(z)}{\zeta_I(\lambda)}$$

and the second term belongs to the ideal generated by ζ_I.
(2) The function R_λ is regular at $z = \lambda$ since it is of the form

$$R_\lambda(z) = \frac{1}{\zeta_I(\lambda)}\,\frac{\zeta_I(z) - \zeta_I(\lambda)}{z - \lambda}\,. \tag{4.301}$$

It is holomorphic in the same strip as $\zeta_I(z)$ and its growth is tempered since this is the case for $\zeta_I(z)$. Using the characterization of J given by Lemma 4.117 (1), one gets that multiplication by R_λ preserves J.

(3) Using (4.300) one gets that the size of $R_\lambda(z)$ on the boundary of a strip at distance > 1 of λ is controlled by

$$|R_\lambda(z)| \leq 1 + \frac{|\zeta_I(z)|}{|\zeta_I(\lambda)|}\,. \tag{4.302}$$

∎

LEMMA 4.121. *Let $N \in \mathbb{N}$ and $Z \subset S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ a subset (with multiplicity) of cardinality N. Then there exists $z \in S^1$ such that*

$$p(z) = \prod_{\lambda \in Z} |z - \lambda| \geq 1\,. \tag{4.303}$$

One has, for $c \geq \log 2$,

$$p(z) \geq 2^N\, e^{-cN}\,. \tag{4.304}$$

except on a set of probability measure $\leq \frac{\log 2}{c}$.

PROOF. One has (cf. [**257**] Lemma 15.17)

$$\frac{1}{2\pi}\int_{S^1} \log|z - \lambda|\,\frac{dz}{z} = 0\,, \tag{4.305}$$

thus the average value of the function $\log p(z)$ is 0 and since it is not constant it must assume a positive value, which gives the first statement.
The function $f(z) = -\log p(z) + N\log 2$ is positive since 2 is the diameter of S^1. Its integral is $N\log 2$ and thus the measure of $E_c = \{z \in S^1 \mid f(z) > cN\}$ is at most $\frac{N\log 2}{cN} = \frac{\log 2}{c}$. On the complement of E_c one has

$$-\log p(z) + N\log 2 \leq cN \quad \text{and} \quad \log p(z) \geq N\log 2 - cN.$$

∎

LEMMA 4.122. *Let $N \in \mathbb{N}$ and $Z \subset \{z \in \mathbb{C} \mid |\Re(z)| \leq 1,\ |\Im(z)| \leq \frac{2}{3}\}$ a subset (with multiplicity) of cardinality N. Then there exists $\alpha \in (-\frac{1}{3}, \frac{1}{3})$ such that*

$$\prod_{\lambda \in Z} |z - \lambda| \geq (3\pi)^{-N}\,, \quad \forall z\,, \quad \Im(z) = \alpha\,. \tag{4.306}$$

PROOF. One has $|z - \lambda| \geq |\Im(z) - \Im(\lambda)|$, thus it is enough to show that there exists $\alpha \in (-\frac{1}{3}, \frac{1}{3})$ such that

$$r(\alpha) = \prod_{s \in \Im(Z)} |\alpha - s| \geq (3\pi)^{-N}\,. \tag{4.307}$$

The rescaled projection $Z' = \frac{3}{2}\Im(Z)$ is a subset (with multiplicity) of $[-1, 1]$ and has cardinality N. By the second part of Lemma 4.121 for $c = 2\log 2$, there exists $\beta \in (-\frac{1}{2}, \frac{1}{2})$ such that

$$\prod_{s \in \frac{3}{2}\Im(Z)} |e^{i\pi\beta} - e^{i\pi s}| \geq 2^{-N}\,.$$

The inequality

$$|e^{i\pi\beta} - e^{i\pi s}| \leq \pi|\beta - s|\,,$$

which follows from $|2\sin(\pi u/2)| \leq \pi|u|$ gives

$$\prod_{s \in \frac{3}{2}\Im(Z)} |\beta - s| \geq (2\pi)^{-N}$$

Taking $\alpha = \frac{2}{3}\beta$ one gets (4.307). $\square$

For each $T \in \mathbb{R}$ we let

$$Z_\zeta(T) = \{\rho \in Z_\zeta \,|\, |\Im(\rho) - T| < 1\}\,. \tag{4.308}$$

By [**239**] Corollary 3.4, one has

$$\operatorname{Card} Z_\zeta(T) = O(\log|T|)\,, \quad |T| \to \infty\,. \tag{4.309}$$

Let then $C < \infty$ be such that

$$\operatorname{Card} Z_\zeta(T) \leq C(1 + \log^+|T|)\,, \quad \forall T \in \mathbb{R}\,. \tag{4.310}$$

LEMMA 4.123. *For any $E \in \mathbb{R}$, $|E| \geq 2$, there exists $T \in \mathbb{R}$, $|T - E| < \frac{1}{3}$ such that*

$$q(z) := \prod_{\rho \in Z_\zeta(T)} |z - \rho| \geq (3\pi)^{-C}\,|T|^{-C\log 3\pi}\,, \tag{4.311}$$

for all z with $\Im(z) = T$ and $-\frac{1}{2} \leq \Re(z) \leq \frac{3}{2}$.

PROOF. Let

$$Z_1 = \{\rho \in Z_\zeta \,|\, |\Im(\rho) - E| \leq \frac{2}{3}\} \quad \text{and} \quad Z = Z_1 - iE - \frac{1}{2}.$$

We let N be the cardinality of Z. By construction Z is a subset (with multiplicity) of $\{z \in \mathbb{C} \,|\, |\Re(z)| \leq 1,\ |\Im(z)| \leq \frac{2}{3}\}$. By Lemma 4.122, there exists $\alpha \in (-\frac{1}{3}, \frac{1}{3})$ such that

$$\prod_{\lambda \in Z} |z - \lambda| \geq (3\pi)^{-N}\,, \quad \forall z\,, \quad \Im(z) = \alpha\,,\ |\Re(z)| \leq 1\,.$$

Let $T = \alpha + E$. One gets

$$\prod_{\rho \in Z_1} |z - \rho| \geq (3\pi)^{-N}\,, \tag{4.312}$$

for all z with $\Im(z) = T$ and $-\frac{1}{2} \leq \Re(z) \leq \frac{3}{2}$.
One has $Z_1 \subset Z_\zeta(T)$ since $|T - E| < \frac{1}{3}$ and hence

$$|\Im(\rho) - E| \leq \frac{2}{3} \Rightarrow |\Im(\rho) - T| < 1\,.$$

Moreover for any element $\rho \in Z_\zeta(T) \setminus Z_1$ one has $|\Im(\rho) - E| > \frac{2}{3}$, hence $|\Im(\rho) - T| \geq \frac{1}{3}$. This, combined with (4.312), gives

$$\prod_{\rho \in Z_\zeta(T)} |z - \rho| \geq (3\pi)^{-\operatorname{Card} Z_\zeta(T)}\,, \tag{4.313}$$

for all z with $\Im(z) = T$ and $-\frac{1}{2} \leq \Re(z) \leq \frac{3}{2}$.
Using (4.310) then gives the required inequality (4.311). $\square$

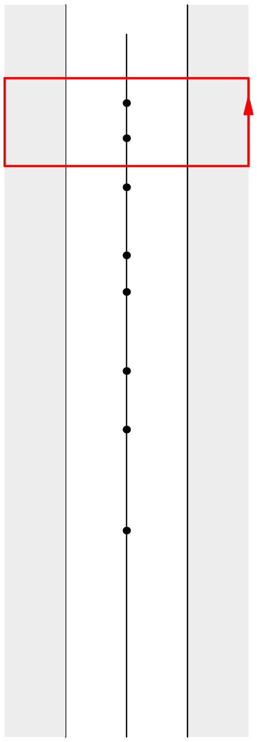


FIGURE 2. Contour γ_N

Now by [**239**] Theorem 3.3, there exists $C' < \infty$ such that for any $T \in \mathbb{R}\,,\ |T| > 1$ one has

$$|\zeta'(z)/\zeta(z) - \sum_{\rho \in Z_\zeta(T)} (z-\rho)^{-1}| \leq C'(1 + \log |T|)\,, \tag{4.314}$$

for all z with $\Im(z) = T$ and $-\frac{1}{2} \leq \Re(z) \leq \frac{3}{2}$.
The absolute convergence of the Euler product gives $C_1 < \infty$ such that

$$|\log|\zeta(3/2 + iT)|| \leq C_1\,, \quad \forall T \in \mathbb{R}\,. \tag{4.315}$$

We then let C_2 be given by

$$C_2 = C \log 2 + 2C' + C_1\,, \tag{4.316}$$

which implies

$$(C \log 2 + 2C')(1 + \log|T|) + C_1 \leq C_2(1 + \log|T|)\,, \tag{4.317}$$

for all T with $|T| > 1$.
We let $C_3 < \infty$ be given by

$$C_3 = (C \log(3\pi) + C_2)\,. \tag{4.318}$$

LEMMA 4.124. *For any $E \in \mathbb{R}$, $|E| \geq 2$, there exists $T \in \mathbb{R}$, with $|T - E| < \frac{1}{3}$, such that*

$$|\zeta(z)| \geq e^{-C_3}\, |T|^{-C_3}\,, \tag{4.319}$$

for all z with $\Im(z) = T$ and $-\frac{1}{2} \leq \Re(z) \leq \frac{3}{2}$.

PROOF. We take T fulfilling the conclusion of Lemma 4.123. Take z with $\Im(z) = T$ and $-\frac{1}{2} \leq \Re(z) \leq \frac{3}{2}$. The horizontal segment $[z, 3/2 + iT]$ does not meet Z_ζ. Integrating the inequality (4.314) along the segment, and taking the real part, gives (with $q(z)$ as in (4.311))

$$|(\Re \log \zeta(z) - \log q(z)) - (\Re \log \zeta(3/2 + iT) - \log q(3/2 + iT))| \\ \leq 2C'(1 + \log |T|)\,. \tag{4.320}$$

Now one has

$$q(3/2 + iT) = \prod_{\rho \in Z_\zeta(T)} |3/2 + iT - \rho|$$

and

$$2^{-\operatorname{Card} Z_\zeta(T)} \leq q(3/2 + iT) \leq 2^{\operatorname{Card} Z_\zeta(T)}$$

which gives

$$|\log q(3/2 + iT)| \leq C \log 2\,(1 + \log |T|)\,. \tag{4.321}$$

It then follows from (4.320) that

$$\begin{aligned} |(\Re \log \zeta(z) - \log q(z))| \leq &\quad (C \log 2 + 2C')(1 + \log |T|) + C_1 \\ \leq &\quad C_2(1 + \log |T|)\,. \end{aligned} \tag{4.322}$$

We thus get using (4.311) that

$$\Re \log \zeta(z) \geq \log q(z) - C_2(1 + \log |T|) \geq -(C \log(3\pi) + C_2)(1 + \log |T|)$$

which gives the required inequality (4.319). □

We now fix a sequence of real numbers T_N, $N \in \mathbb{Z}$, $|N| \geq 1$, such that

(1) For all N one has $|T_N - N| < \frac{1}{3}$.
(2) The inequality (4.319) holds for all T_N.

We then define operators $\Pi(N)$ as

$$\Pi(N) = \frac{1}{2\pi i} \int_{\gamma_N} R_\lambda \, d\lambda\,, \tag{4.323}$$

where γ_N, for $N \geq 1$, is the rectangular contour (cf. Figure 2) with vertices at the four points

$$\{-\frac{1}{2} + iT_N\,,\ \frac{3}{2} + iT_N\,,\ \frac{3}{2} + iT_{N+1}\,,\ -\frac{1}{2} + iT_{N+1}\}\,.$$

For $N \leq -1$ we let γ_N be the rectangular contour with vertices at the four points

$$\{-\frac{1}{2} + iT_{N-1}\,,\ \frac{3}{2} + iT_{N-1}\,,\ \frac{3}{2} + iT_N\,,\ -\frac{1}{2} + iT_N\}\,,$$

while γ_0 has vertices

$$\{-\frac{1}{2} + iT_{-1}\,,\ \frac{3}{2} + iT_{-1}\,,\ \frac{3}{2} + iT_1\,,\ -\frac{1}{2} + iT_1\}\,.$$

LEMMA 4.125. *The operators $\Pi(N)$ acting on H^1 have the following properties.*

(1) *Each $\Pi(N)$ is an idempotent: $\Pi(N)^2 = \Pi(N)$.*
(2) *The sequence $\Pi(N)$ is of tempered growth.*
(3) *The rank of $\Pi(N)$ is $O(\log |N|)$ for $|N| \to \infty$.*

(4) *For any $f \in \mathcal{S}(I)$ the sequence $f\,\Pi(N)$ is of rapid decay and*

$$\sum_{\mathbb{Z}} f\,\Pi(N) = f \mod J\,. \tag{4.324}$$

PROOF. (1) By Lemma 4.120 the multiplication by R_λ agrees with the resolvent $\frac{1}{\lambda - z}$ modulo the ideal J generated by ζ_I. In particular, one gets the identity modulo J

$$R_\lambda\, R_{\lambda'} \sim (\lambda' - \lambda)^{-1}(R_\lambda - R_{\lambda'})\,,\quad \forall\lambda\,,\ \lambda' \notin Z\,, \tag{4.325}$$

using

$$(z-\lambda)P(z-\lambda') \in J\,,\quad P = R_\lambda\, R_{\lambda'} - (\lambda'-\lambda)^{-1}(R_\lambda - R_{\lambda'})\,,$$

so that $P \in J$, using $P \sim R_\lambda(z-\lambda)P(z-\lambda')R_{\lambda'}$ modulo J.
Since R_λ is a holomorphic function of $\lambda \notin Z$, one can replace the contour γ_N used in (4.323) by a slightly smaller concentric rectangular path. Then

$$\begin{aligned} \Pi(N)^2 = &\ \tfrac{1}{(2\pi i)^2} \int_{\gamma_N}\int_{\gamma'_N} R_\lambda\, R_{\lambda'}\, d\lambda d\lambda' \\ \sim &\ \tfrac{1}{(2\pi i)^2} \int_{\gamma_N}\int_{\gamma'_N} (\lambda'-\lambda)^{-1}(R_\lambda - R_{\lambda'})\, d\lambda d\lambda'\,, \end{aligned}$$

which gives $\Pi(N)$, using

$$\int_{\gamma'_N} (\lambda'-\lambda)^{-1}\, d\lambda' = 0\,,\quad \int_{\gamma_N} (\lambda'-\lambda)^{-1}\, d\lambda = -2\pi i\,.$$

(2) By Lemma 4.120, the size of R_λ is controlled by $\frac{1}{|\zeta_I(\lambda)|}$, so that the size of $\Pi(N)$ is controlled by $\sup_{\gamma_N} \frac{1}{|\zeta_I(\lambda)|}$. Using (4.297) one can, in order to show that this sequence is of tempered growth, replace ζ_I by ζ so that we need to estimate $\sup_{\gamma_N} \frac{1}{|\zeta(\lambda)|}$. Thus we get the result, using the inequality (4.319) which holds for the horizontal part of γ_N. On the vertical part one uses the bound (4.315) for $\lambda = \frac{3}{2} + is$ and

$$|\zeta(-\frac{1}{2}+is)| = |\zeta(-\frac{1}{2}-is)| = |\zeta(\frac{3}{2}+is)|\pi^{-1}|\Gamma(\frac{3}{4}+i\frac{s}{2})/\Gamma(-\frac{1}{4}-i\frac{s}{2})|,$$

which follows from the functional equation. Since

$$|\Gamma(\frac{3}{4}+i\frac{s}{2})/\Gamma(-\frac{1}{4}-i\frac{s}{2})| = |\Gamma(\frac{3}{4}+i\frac{s}{2})/\Gamma(-\frac{1}{4}+i\frac{s}{2})| = |-\frac{1}{4}+i\frac{s}{2}|\,,$$

one gets that $\sup_{\gamma_N} \frac{1}{|\zeta(\lambda)|}$ has tempered growth and the same holds for $\sup_{\gamma_N} \frac{1}{|\zeta_I(\lambda)|}$ and for $\Pi(N)$.
(3) The equality (4.323) defines a function $\Pi(N)(z)$. Let $z_0 \in Z_\zeta$ be a zero of zeta of order $n > 0$. The Taylor expansion of $\zeta_I(z)$ at $z = z_0$ is $O(z-z_0)^n$. Thus (4.300) shows that, for $\lambda \in \gamma_N$, the Taylor expansion of $R_\lambda(z)$ at $z = z_0$ agrees with that of $(\lambda - z)^{-1}$ up to order n, i.e.

$$R_\lambda(z) - \frac{1}{\lambda - z} = O(z-z_0)^n\,.$$

Thus one gets

$$\Pi(N)(z) - \frac{1}{2\pi i}\int_{\gamma_N} \frac{1}{\lambda - z}\, d\lambda = O(z-z_0)^n\,, \tag{4.326}$$

while

$$\frac{1}{2\pi i}\int_{\gamma_N}\frac{1}{\lambda - z}\,d\lambda = \begin{cases} 1 & \text{for } z \text{ inside } \gamma_N, \\ 0 & \text{for } z \text{ outside } \gamma_N. \end{cases} \tag{4.327}$$

Thus, by Corollary 4.118, we get that in H^1 the operator $\Pi(N)$ is the projection on the subset of zeros inside γ_N. Since the cardinality of this set (with multiplicity) is $O(\log N)$ one gets the desired result.

(4) The equality (4.326) is unaltered if we replace the function $\zeta_I(z)$ in the construction of R_λ by a function $h \in J$ with the same zeros. Thus we obtain the same operator $\Pi(N)$ on H^1 after the replacement $\zeta_I \mapsto h$. For each N we take

$$h_N(z) = e^{(z-iN)^2}\,\zeta_I(z)$$

and use h_N instead of ζ_I so that the formula (4.323) becomes

$$\Pi(N) = \frac{1}{2\pi i}\int_{\gamma_N}\frac{1}{\lambda - z}\left(1 - \frac{h_N(z)}{h_N(\lambda)}\right)d\lambda\,. \tag{4.328}$$

Note that the lower bound on $|\zeta_I(\lambda)|$ for $\lambda \in \gamma_N$ gives a similar lower bound on $h_N(\lambda)$ since $|\lambda - iN| \leq 4$ for $\lambda \in \gamma_N$. Moreover, for z outside of the rectangle bounded by γ_N, the integral of $\frac{d\lambda}{\lambda - z}$ vanishes and (4.328) gives

$$\Pi(N) = -\frac{1}{2\pi i}\int_{\gamma_N}\frac{h_N(z)}{(\lambda - z)h_N(\lambda)}d\lambda\,.$$

We thus get an estimate of the form

$$|\Pi(N)(z)| \leq C\,N^k(1+|z|)^k\,e^{-(\Im(z)-N)^2}\,,\quad \forall z \in I\,, \tag{4.329}$$

where the constant $C < \infty$ and the integer k are independent of N. Let then $f \in \mathcal{S}(I)$. One has, for any m, a constant C_m such that $|f(z)| \leq C_m\,(1+|z|)^{-m}$, $\forall z \in I$. We thus get

$$|f(z)\,\Pi(N)(z)| \leq C\,C_m\,N^k\,(1+|z|)^{k-m}\,e^{-(\Im(z)-N)^2}\,,\quad \forall z \in I\,.$$

Using, for $q > 0$, the inequality

$$\sup_{z\in I}\,(1+|z|)^{-q}\,e^{-(\Im(z)-N)^2} \leq (1+N/2)^{-q}\,,\quad \forall N \geq 2q\,,$$

one gets that the sequence $p(f\,\Pi(N))$ is of rapid decay for any of the seminorms p defining the topology of $\mathcal{S}(I)$, which can be taken explicitly of the form

$$p_{k,m}(f) = \sum_0^m \frac{1}{n!}\,\sup_I(1+|z|)^k|\partial^n f(z)|\,. \tag{4.330}$$

The equality (4.324), valid modulo J, then follows from Corollary 4.118. ☐

Proof. of Theorem 4.116. The seminorms (4.330) are submultiplicative, i.e.

$$p_{k,m}(fg) \leq p_{k,m}(f)\,p_{k,m}(g)\,,\quad \forall f\,,\ g\,. \tag{4.331}$$

Fixing (k,m) we get that the operator of multiplication by f in H^1 is the convergent sum of the operators of multiplication by $f\Pi(N)$, where the sequence of rapid decay $p_{k,m}(f\,\Pi(N))$ controls the operator norm of $f\Pi(N)$. The existence of Auerbach basis in finite dimensional Banach spaces implies the inequality $N_1(A) \leq \mathrm{rank}A\|A\|$ for the nuclear norm $N_1(A)$ of a finite rank operator A in a Banach space in terms of its operator norm $\|A\|$. Thus part 3 of Lemma 4.125 then shows that the sum

(4.324) converges in the nuclear norm. It follows, writing f as the product of two elements of $\mathcal{S}(I)$, that the operator of multiplication by f is a binuclear operator (a product of two nuclear operators) so that its trace is well defined. The trace is obtained as the sum of the traces of the finite rank operators $f\,\Pi(N)$. For each N the evaluation map (4.299) gives a surjection from $\mathcal{S}(I)$ to

$$\oplus\,\mathbb{C}^{(n_\rho)}\,,\quad \rho\in Z_\zeta\,,\quad \rho\ \text{ inside }\gamma_N$$

where, as above, n_ρ is the multiplicity of ρ and $\mathbb{C}^{(n_\rho)}\sim\mathbb{C}^{n_\rho}$ is the space of dimension n_ρ of jets of order n_ρ. The trace of the operator of multiplication by f acting on jets of order n_ρ is $n_\rho\,f(\rho)$ since the matrix of the operator is triangular in the basis $(z-\rho)^k$, with diagonal entries equal to $f(\rho)$. Thus the trace of the finite rank operators $f\,\Pi(N)$ is

$$\mathrm{Tr}(f\,\Pi(N))=\sum_{\rho\,\text{inside}\,\gamma_N} n_\rho\,f(\rho)\,.$$

This shows that for any of the seminorms $p_{k,m}$ the operator of multiplication by f is a binuclear operator with trace given by

$$\mathrm{Tr}(f)=\sum_{\rho\in Z_\zeta} n_\rho\,f(\rho)\,.$$

Applying this result to the Fourier transform $\hat{f}(1,z)$ gives the expected formula (4.296). ◻

REMARK 4.126. The map $\mathfrak{E}$ considered in Chapter 2 involves a slightly different normalization than the map $\mathbf{\Sigma}$ considered here. This depended on the fact that, in working in the Hilbert space context, the result of the summation on $\mathbb{K}^*$ was then multiplied by $|x|^{1/2}$ (cf. [**71**]).
This introduces a shift of $\frac{1}{2}$ in the spectrum, using the following relation between the Fourier transforms

$$\widehat{\mathfrak{E}(f)}(\tilde{\chi},\rho)=\widehat{\mathbf{\Sigma}(f)}(\tilde{\chi},\rho+\frac{1}{2}) \tag{4.332}$$

where as above

$$\hat{h}(\tilde{\chi},\rho)=\int_{C_{\mathbb{K}}} h(u)\tilde{\chi}(u)\,|u|^\rho\,d^*u\,. \tag{4.333}$$

One has $\mathfrak{E}=\Delta^{1/2}\mathbf{\Sigma}$ where $\Delta^z(f)(g)=|g|^z f(g)$ is defined for any $z\in\mathbb{C}$ (cf. (4.350) below) so that $\Delta^{1/2}$ relates the two different normalizations. Note that the representations $\vartheta_{\mathrm{m}}(\gamma)$ are then related by

$$\Delta^{-1/2}\vartheta_{\mathrm{m}}^{\mathfrak{E}}(\gamma)\Delta^{1/2}=|\gamma|^{-1/2}\vartheta_{\mathrm{m}}^{\mathbf{\Sigma}}(\gamma)\,, \tag{4.334}$$

so that, at the formal level, the trace computation of Corollary 2.48 corresponds to that of Theorem 4.116. Indeed, from (4.334) we have

$$\vartheta_{\mathrm{m}}^{\mathbf{\Sigma}}(f)=\int_{C_{\mathbb{K}}} f(\gamma)\,\vartheta_{\mathrm{m}}^{\mathbf{\Sigma}}(\gamma)\,d^*\gamma=\Delta^{-1/2}\int_{C_{\mathbb{K}}} h(\gamma)\,\vartheta_{\mathrm{m}}^{\mathfrak{E}}(\gamma)\,d^*\gamma\;\Delta^{1/2}\,, \tag{4.335}$$

where

$$h(\gamma)=\,|\gamma|^{1/2}\,f(\gamma). \tag{4.336}$$

Thus, an equality of the form

$$\operatorname{Tr} \underline{\vartheta}_{\mathrm{m}}^{\mathfrak{C}}(h) = \sum_{\rho \in \mathbb{C}/N^{\perp} \mid L(\tilde{\chi}, \frac{1}{2}+\rho)=0} \hat{h}(\tilde{\chi}, \rho) \tag{4.337}$$

corresponds to (4.296). Indeed, the Fourier transform $\hat{h}(\tilde{\chi}, \rho)$ satisfies

$$\begin{aligned} \hat{h}(\tilde{\chi}, \rho) &= \int_{C_{\mathbb{K}}} h(u)\tilde{\chi}(u)\,|u|^{\rho}\, d^*u \\ &= \int_{C_{\mathbb{K}}} f(u)\tilde{\chi}(u)\,|u|^{\rho+1/2}\, d^*u \\ &= \hat{f}(\tilde{\chi}, 1/2+\rho), \end{aligned} \tag{4.338}$$

where h and f are related as in (4.336). Thus, the shift by $1/2$ in (4.337) is absorbed in (4.338) and this gives (4.296).

5.6. Trace pairing and vanishing.

The commutativity of the convolution product implies the following vanishing result.

LEMMA 4.127. *Suppose given an element $f \in \mathcal{V} \subset \mathbf{S}(C_{\mathbb{K}})$, where $\mathcal{V}$ is the range of the reduction map as in Definition 4.114. Then one has*

$$\underline{\vartheta}_{\mathrm{m}}(f)|_{H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})} = 0. \tag{4.339}$$

PROOF. By Lemma 4.115, convolution by $\mathbf{S}(C_{\mathbb{K}})$ leaves $\mathcal{V}$ invariant, i.e. $\mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$ is stable under convolution by $\mathbf{S}(C_{\mathbb{K}})$. The result follows by showing that, for $f \in \mathcal{V}$, the operator $\underline{\vartheta}_{\mathrm{m}}(f)$ maps any element $\xi \in \mathbf{S}(C_{\mathbb{K}})$ to an element in $\mathcal{V}$, and therefore the induced map on the quotient of $\mathbf{S}(C_{\mathbb{K}})$ by $\mathcal{V}$ is trivial. Since $\mathcal{V}$ is a submodule of $\mathbf{S}(C_{\mathbb{K}})$ for the action of $\mathbf{S}(C_{\mathbb{K}})$ by convolution we obtain

$$\underline{\vartheta}_{\mathrm{m}}(f)\xi = f \star \xi = \xi \star f \in \mathcal{V},$$

where $\star$ is the convolution product of (4.294). ☐

This makes it possible to define a trace pairing as follows.

REMARK 4.128. *The pairing*

$$f_1 \otimes f_2 \mapsto \langle f_1, f_2 \rangle_{H^1} := \operatorname{Tr}(\underline{\vartheta}_{\mathrm{m}}(f_1 \star f_2)|H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})) \tag{4.340}$$

descends to a well-defined pairing on $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}}) \otimes H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$.

5.7. Weil's explicit formula as a trace formula.

The fact that we do not work here with Hilbert spaces means, in particular, that we do not have the restriction of unitarity in the setting of Theorem 4.116 Thus, we will be able to obtain a trace formula which is not only semilocal as in [**71**]. The RH will then be equivalent to a positivity statement. This is similar to what happens in [**224**], except that here we formulate the trace formula in a cohomological version, which is closer to what happens in the classical setting with the action of Frobenius on étale cohomology.

We first introduce some preliminary notation. For a local field $K = \mathbb{K}_v$, one can choose an additive character e_K, in the following way:

- $e_{\mathbb{R}}(x) = e^{-2\pi i x}$, for all $x \in \mathbb{R}$,
- $e_{\mathbb{C}}(z) = e^{-2\pi i(z+\bar{z})}$, for all $z \in \mathbb{C}$
- If K is a non-Archimedean local field with maximal compact subring $\mathcal{O}$, then the character e_K fulfills $\operatorname{Ker} e_K = \mathcal{O}$.

Suppose given a non-trivial character α of the additive group $\mathbb{A}_{\mathbb{K}}$, which is trivial on $\mathbb{K} \subset \mathbb{A}_{\mathbb{K}}$. We write

$$\alpha = \prod_{v \in \Sigma_{\mathbb{K}}} \alpha_v, \tag{4.341}$$

for the decomposition of α as a product of its restrictions to the local fields $\alpha_v = \alpha|_{\mathbb{K}_v}$.
The notion of a *differental idèle* (cf. [**293**] Def. VII-2-4) extends the notion of the canonical class for an algebraic curve C, from the setting of function fields $\mathbb{F}_q(C)$ to arbitrary global fields $\mathbb{K}$ (cf. also §6 below for the role of the canonical class in the function fields case).

Definition 4.129. *An idèle $a = (a_v) \in \mathbb{A}_{\mathbb{K}}^*$ is called a* differental idèle *if, for every place $v \in \Sigma_{\mathbb{K}}$, it satisfies*

$$\alpha_v(x) = e_{\mathbb{K}_v}(a_v x), \quad \forall x \in \mathbb{K}_v \tag{4.342}$$

In other words (4.342) means that we have

$$\alpha(x) = (\prod_{v \in \Sigma_{\mathbb{K}}} e_{\mathbb{K}_v})(ax), \quad \forall x \in \mathbb{A}_{\mathbb{K}}. \tag{4.343}$$

We recall also (cf. [**293**] Proposition VII-2-6) that a differental idèle $a = (a_v)$ satisfies

$$|a| = q^{2-2g} \qquad \text{or} \qquad |a| = D^{-1}, \tag{4.344}$$

for a function field $\mathbb{F}_q(C)$ and for a number field, respectively, where in the number field case D denotes the discriminant of the field (which can be thought of as the volume of a fundamental region for the ring of integers $\mathcal{O}_{\mathbb{K}}$ of $\mathbb{K}$).

We can then write the trace formula in the following way, [**75**], which gives the *geometric side* of the trace formula.

Theorem 4.130. *For $f \in \mathbf{S}(C_{\mathbb{K}})$ let $\underline{\vartheta}_{\mathrm{m}}(f)$ be the operator* (4.293) *acting on the space $H^1 = H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$. Then the trace is given by*

$$\begin{aligned} \operatorname{Tr}(\underline{\vartheta}_{\mathrm{m}}(f)|H^1) = \; & \hat{f}(0) + \hat{f}(1) - (\log|a|)\, f(1) \\ & - \sum_{v \in \Sigma_{\mathbb{K}}} \int'_{(\mathbb{K}_v^*, e_{\mathbb{K}_v})} \frac{f(u^{-1})}{|1-u|}\, d^*u. \end{aligned} \tag{4.345}$$

Proof. Using the Riemann-Weil explicit formula (2.348) of Chapter 2, together with Theorem 4.116, one gets

$$\operatorname{Tr}(\underline{\vartheta}_{\mathrm{m}}(f)|H^1) = \hat{f}(0) + \hat{f}(1) - \sum_{v \in \mathbb{K}_v} \int'_{(\mathbb{K}_v^*, \alpha_v)} \frac{f(u^{-1})}{|1-u|}\, d^*u, \tag{4.346}$$

where one uses the global character α to fix the local normalizations of the principal values in the last term of the formula.

The fact that the terms

$$\int'_{(\mathbb{K}_v^*,\alpha_v)} \frac{f(u^{-1})}{|1-u|}\, d^*u \tag{4.347}$$

depend on the global character α needs to be corrected if we want to really have "local" contributions that only use the information coming from the local field $\mathbb{K}_v$. This can be done via the following lemma, using differental idèles.

LEMMA 4.131. *The principal values* (4.347) *can be computed in the form*

$$\int'_{(\mathbb{K}_v^*,\alpha_v)} \frac{f(u^{-1})}{|1-u|}\, d^*u = (\log|a_v|)\, f(1) + \int'_{(\mathbb{K}_v^*,e_{\mathbb{K}_v})} \frac{f(u^{-1})}{|1-u|}\, d^*u. \tag{4.348}$$

PROOF. This follows from Proposition 2.11, (3) and (2.100) of Chapter 2. □

This completes the proof of Theorem 4.130. □

REMARK 4.132. It is quite important that Theorem 4.130 can be derived directly without using the Riemann–Weil explicit formula, following [**71**]. This was done by Ralf Meyer in [**224**].

5.8. Weil positivity and the Riemann Hypothesis.

We introduce an involution for elements $f \in \mathbf{S}(C_{\mathbb{K}})$ by setting

$$f^*(g) = \overline{f(g^{-1})}. \tag{4.349}$$

We also consider a one-parameter group $z \mapsto \Delta^z$ of automorphisms of the convolution algebra $\mathbf{S}(C_{\mathbb{K}})$, with the convolution product (4.294), by setting

$$\Delta^z(f)(g) = |g|^z\, f(g), \tag{4.350}$$

for $f \in \mathbf{S}(C_{\mathbb{K}})$ and $z \in \mathbb{C}$. Since (4.350) is given by multiplication by a character, it satisfies

$$\Delta^z(f \star h) = \Delta^z(f) \star \Delta^z(f), \quad \forall f, h \in \mathbf{S}(C_{\mathbb{K}}). \tag{4.351}$$

We consider also the involution

$$f \mapsto f^\sharp = \Delta^{-1} f^*, \quad \text{with} \quad f^\sharp(g) = |g|^{-1}\overline{f(g^{-1})}. \tag{4.352}$$

The reformulation, originally due to A. Weil (cf. [**294**]), of the Riemann Hypothesis in our setting is given by the following statement.

PROPOSITION 4.133. *The following two conditions are equivalent:*

- *All L-functions with Grössencharakter on $\mathbb{K}$ satisfy the Riemann Hypothesis.*
- *The trace pairing* (4.340) *satisfies the positivity condition*

$$\langle \Delta^{-1/2} f, \Delta^{-1/2} f^* \rangle \geq 0, \quad \forall f \in \mathbf{S}(C_{\mathbb{K}}). \tag{4.353}$$

PROOF. Let $W(\gamma) = |\gamma|^{-1/2}\, \underline{\vartheta}_{\mathrm{m}}(\gamma)$. Then, RH for L-functions with Grössencharakter on $\mathbb{K}$ is equivalent, by [**294**], to the positivity

$$\mathrm{Tr}(W(f \star f^*)) \geq 0, \quad \forall f \in \mathbf{S}(C_{\mathbb{K}}). \tag{4.354}$$

Thus, in terms of the representation $\underline{\vartheta}_{\mathrm{m}}$ we are considering here we have $W(f) = \underline{\vartheta}_{\mathrm{m}}(\Delta^{-1/2} f)$ and using the multiplicative property (4.351) of Δ^z we rewrite (4.354) in the equivalent form (4.353). It is worth sketching the argument showing the equivalence. Consider an algebra $\mathcal{B}$ of holomorphic functions, endowed with the pointwise product, in a domain $\Omega \subset \mathbb{C}$ with $\bar{\Omega} = \Omega$. Assume that $\mathcal{B}$ is stable under the involution

$$f^*(z) := \overline{f(\bar{z})}\,, \quad \forall z \in \Omega\,. \tag{4.355}$$

Any element $z \in \Omega$ defines a self-adjoint linear form on $\mathcal{B}$ by

$$L_z(f) = f(z) + \overline{f(\bar{z})}\,, \quad \forall f \in \mathcal{B}\,. \tag{4.356}$$

This linear form is positive, i.e. it fulfills $L(ff^*) \geq 0$ for all $f \in \mathcal{B}$ as soon as $z \in \Omega \cap \mathbb{R}$. Indeed in that case one gets

$$L_z(ff^*) = 2f(z)\,\overline{f(\bar{z})} = 2|f(z)|^2 \geq 0\,. \tag{4.357}$$

This is enough to check the positivity (4.353) assuming RH. We refer to [**294**] for the converse. ☐

In terms of the involution (4.352) we can reformulate Lemma 4.133 in the following equivalent way.

Corollary 4.134. *The following conditions are equivalent*

- *All L-functions with Grössencharakter on* $\mathbb{K}$ *satisfy the Riemann Hypothesis.*
- *The trace pairing* (4.340) *satisfies* $\langle f, f^\sharp\rangle \geq 0$, *for all* $f \in \mathbf{S}(C_{\mathbb{K}})$.

Proof. In (4.353) we write $\Delta^{-1/2} f = h$. This gives

$$\Delta^{-1/2} f^* = \Delta^{-1/2}(\Delta^{1/2} h)^* = \Delta^{-1} h^* = h^\sharp$$

and the result follows, since $\Delta^{-1/2}$ is an automorphism of $\mathbf{S}(C_{\mathbb{K}})$. ☐

The vanishing result of Lemma 4.127, for elements in the range $\mathcal{V} \subset \mathbf{S}(C_{\mathbb{K}})$ of the reduction map $\mathrm{Tr} \circ \pi$ from adèles, then gives the following.

Proposition 4.135. *The elements* $f \star f^\sharp$ *considered in Corollary 4.134 above have the following properties.*

1. *The trace pairing* $\langle f, f^\sharp\rangle$ *vanishes for all* $f \in \mathcal{V}$, *i.e. when* f *is the restriction* $\mathrm{Tr} \circ \pi$ *of an element of* $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$.
2. *By adding elements of* $\mathcal{V}$ *one can make the values*

$$f \star f^\sharp(1) = \int_{C_{\mathbb{K}}} |f(g)|^2\, |g|\, d^*g < \epsilon \tag{4.358}$$

for arbitrarily small $\epsilon > 0$.

Proof. (1) The vanishing result of Lemma 4.127 shows that

$$\underline{\vartheta}_{\mathrm{m}}(f)|_{H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})} = 0$$

for all $f \in \mathcal{V}$. Thus, the trace pairing satisfies $\langle f, h\rangle = 0$, for $f \in \mathcal{V}$ and for all $h \in \mathbf{S}(C_{\mathbb{K}})$. In particular this applies to the case $h = f^\sharp$.

(2) This follows from the surjectivity of the map $\mathfrak{E}$ for the value $\delta = 0$ (cf. Appendix 1 of [**71**]). ☐

Proposition 4.135 shows that the trace pairing admits a large radical given by all functions that extend to adèles. Thus, one is led in all cases to divide by this radical, hence to work with the cohomology $H^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ described above.

We now proceed to compare the setting we have developed in terms of the noncommutative geometry of the adèle class space with the classical algebro-geometric setting of the Weil proof of RH for function fields.

6. The Weil proof for function fields

We review the Weil proof of the Riemann Hypothesis for function fields, with the idea of gradually building a dictionary that will make it possible to rephrase some key notions in terms of the noncommutative geometry of the adèle class space and transfer them to the number field case in this way.

6.1. Function fields and their zeta functions.

Let $\mathbb{K}$ be a global field of positive characteristic $p > 0$. Then there exists a smooth projective curve over a finite field $\mathbb{F}_q$, with $q = p^r$ for some $r \in \mathbb{N}$, such that

$$\mathbb{K} = \mathbb{F}_q(C) \tag{4.359}$$

is the field of functions of C.

As we discussed in the section on motives in Chapter 1, the curve C over $\mathbb{F}_q$ has a zeta function of the form

$$Z_C(T) = \exp\left(\sum_{n=1}^{\infty} \frac{\#C(\mathbb{F}_{q^n})}{n} T^n\right), \tag{4.360}$$

with $\log Z_C(T)$ the generating function for the number of points of C over the fields $\mathbb{F}_{q^n}$. One writes

$$\zeta_C(s) = Z_C(q^{-s}). \tag{4.361}$$

We denote by $\Sigma_{\mathbb{K}}$ the set of places of $\mathbb{K}$. A place $v \in \Sigma_{\mathbb{K}}$ is a Galois orbit of points of $C(\bar{\mathbb{F}}_q)$. The degree $n_v = \deg(v)$ is its cardinality, namely the number of points in the orbit of the Frobenius acting on the fiber of the natural map from points to places

$$C(\bar{\mathbb{F}}_q) \to \Sigma_{\mathbb{K}}. \tag{4.362}$$

This means that there are n_v conjugate points defined over $\mathbb{F}_{q^{n_v}}$, the residue field of the local field $\mathbb{K}_v$.

The zeta function of the function field $\mathbb{K}$ is just the zeta function of C considered above

$$\zeta_{\mathbb{K}}(s) = \zeta_C(s), \tag{4.363}$$

which converges for $\Re(s) > 1$.

Lemma 4.136. *The zeta function* (4.363) *has an Euler product expression of the form*

$$\zeta_{\mathbb{K}}(s) = \prod_{v \in \Sigma_{\mathbb{K}}} (1 - q^{-n_v s})^{-1}. \tag{4.364}$$

PROOF. One writes

$$\log \prod_v (1-q^{-n_v s})^{-1} = \sum_v \sum_{m=1}^{\infty} \frac{q^{-sn_v m}}{m}$$
$$= \sum_{n=1}^{\infty} q^{-sn} \sum_{n_v|n} \frac{n_v}{n} = \sum_{n=1}^{\infty} \left(\sum_{n_v|n} n_v \right) \frac{q^{-sn}}{n}.$$

The expression $\sum_{n_v|n} n_v$ is in fact precisely $\#C(\mathbb{F}_{q^n})$ so that one gets (4.364). $\square$

It is useful to consider also an equivalent formulation in terms of divisors on C. Let $\mathrm{Div}(C)$ denote the abelian group of divisors on C, i.e. finite integer combinations of points. For a point P we let $\deg(P)$ be the degree of the residue field at P and for a divisor $D = \sum_P n_P P$ we let $\deg(D) = \sum_P n_P \deg(P)$. We say that a divisor $D = \sum_P n_P P$ is *effective*, or that $D \geq 0$, if it has $n_p \geq 0$ for all P. One defines the norm of a divisor D as

$$N(D) = q^{\deg(D)}. \tag{4.365}$$

LEMMA 4.137. *For* $\Re(s) > 1$ *the zeta function* (4.363) *can be written equivalently as*

$$\zeta_{\mathbb{K}}(s) = \sum_{D \geq 0} N(D)^{-s}. \tag{4.366}$$

PROOF. One can write $N(D) = q^{\deg(D)} = \prod_P N(P)^{n_P}$. This gives

$$\prod_P \sum_{n_P=0}^{\infty} N(P)^{-n_P s} = \sum_{D \geq 0} N(D)^{-s},$$

since for $\Re(s) > 1$ one has absolute convergence. By Lemma 4.136, the left hand side gives $\zeta_C(s) = \prod_P (1 - N(P)^{-s})^{-1}$. $\square$

For a divisor D on C one can consider the space

$$H^0(D) = \{ f \in \mathbb{K}^* \mid \mathrm{div}(f) + D \geq 0 \} \cup \{0\}. \tag{4.367}$$

One denotes by $\ell(D)$ the dimension of $H^0(D)$. One then has

$$\# H^0(D) = q^{\ell(D)}.$$

The Riemann–Roch formula formula for the curve C states that

$$\ell(D) - \ell(\kappa_C - D) = \deg(D) - g + 1, \tag{4.368}$$

where κ_C is the canonical divisor on C, which has degree $\deg(\kappa_C) = 2g-2$, and $h^0(\kappa_C) = g$.

Notice that, by (4.367), if $\deg(D) > 2g-2$ then $\ell(\kappa_C - D) = 0$ so that by (4.368) one has $\ell(D) = \deg(D) - g + 1$.

Divisors can be considered up to the equivalence relation given by adding a *principal divisor*,

$$D \sim D' \iff D - D' = \mathrm{div}(f), \tag{4.369}$$

for some $f \in \mathbb{K}^*$. The Picard group is the quotient $\mathrm{Pic}(C) = \mathrm{Div}(C)/\sim$. One denotes by $\mathrm{Pic}^n(C)$ the classes $[D] \in \mathrm{Pic}(C)$ with $\deg[D] = n$.

Notice that both $\deg(D)$ and $N(D)$ are well defined on the equivalence classes. We write $\deg[D]$ and $N[D]$ for the resulting functions on $\mathrm{Pic}(C)$. We then can write the zeta function as

$$\begin{aligned}\zeta_{\mathbb{K}}(s) = & \textstyle\sum_{n=0}^{\infty} \#\{D \geq 0,\, \deg(D) = n\}\, q^{-ns} \\ = & \textstyle\sum_{[D]\in \mathrm{Pic}(C)} \#\{D' \in [D],\, D' \geq 0\}\, N(D)^{-s}.\end{aligned}$$

The counting of $\{D' \in [D],\, D' \geq 0\}$ corresponds to the counting of points in the projectivized $\mathbb{P}H^0(D)$, hence one obtains

$$\begin{aligned}\zeta_{\mathbb{K}}(s) = & \textstyle\sum_{[D]\in \mathrm{Pic}(C)} \frac{q^{\ell(D)}-1}{q-1}\; q^{-s\deg[D]} \\ = & \textstyle\sum_{n=0}^{\infty} \sum_{[D]\in \mathrm{Pic}^n(C)} \frac{q^{\ell(D)}-1}{q-1}\; q^{-sn}.\end{aligned} \tag{4.370}$$

Using in (4.370) the Riemann–Roch formula (4.368) and the fact that, for $n > 2g-2$, one has $\ell(D) = \deg[D]-g+1$, one obtains that the zeta function $\zeta_{\mathbb{K}}(s)$ can be written as a product

$$Z_C(T) = \frac{P(T)}{(1-T)(1-qT)}, \tag{4.371}$$

where $P(T)$ is a polynomial of degree $2g$ and integer coefficients

$$P(T) = \prod_{j=1}^{2g} (1-\lambda_j T). \tag{4.372}$$

Equating (4.371) and (4.360) one obtains

$$\#C(\mathbb{F}_{q^n}) = q^n + 1 - \sum_{j=1}^{2g} \lambda_j. \tag{4.373}$$

The Riemann–Roch formula (4.368) also implies that the zeta function $\zeta_{\mathbb{K}}(s)$ satisfies the functional equation

$$\zeta_{\mathbb{K}}(1-s) = q^{(1-g)(1-2s)}\zeta_{\mathbb{K}}(s). \tag{4.374}$$

Another important reformulation of the zeta function can be given in terms of étale cohomology. Let $\bar{C} = C \times_{\mathbb{F}_q} \bar{\mathbb{F}}_q$, and let ℓ be a prime $\ell \neq p$. Then one can consider the étale cohomology groups $H^i(\bar{C}, \mathbb{Q}_\ell)$, for $i = 0, 1, 2$. The Frobenius

$$\mathrm{Fr} : \bar{C} \to \bar{C} \tag{4.375}$$

has the property that $\mathrm{Fr}(P) = P$ iff $P \in C(\mathbb{F}_q)$ and $\mathrm{Fr}^n(P) = P$ iff $P \in C(\mathbb{F}_{q^n})$. Thus, the coefficients $\#C(\mathbb{F}_{q^n})$ that appear in the zeta function can be rewritten as

$$\#C(\mathbb{F}_{q^n}) = \#\mathrm{Fix}(\mathrm{Fr}^n : \bar{C} \to \bar{C}). \tag{4.376}$$

The *Lefschetz fixed point formula* for étale cohomology then shows that

$$\#C(\mathbb{F}_{q^n}) = \sum_{i=0}^{2} (-1)^i \mathrm{Tr}\left(\mathrm{Fr}^n |\, H^i(\bar{C}, \mathbb{Q}_\ell)\right). \tag{4.377}$$

Thus, the zeta function can be written in the form

$$\zeta_{\mathbb{K}}(s) = \prod_{i=0}^{2} \left(\exp\left(\sum_{n=1}^{\infty} \mathrm{Tr}(\mathrm{Fr}^n |\, H^i(\bar{C}, \mathbb{Q}_\ell)) \frac{q^{-sn}}{n} \right) \right)^{(-1)^i}. \tag{4.378}$$

Equivalently, one writes

$$Z_C(T) = \frac{P_1(T)}{P_0(T)P_2(T)}, \quad \text{with} \quad P_i(T) = \det(1 - \mathrm{Fr}\, T |\, H^i(\bar{C}, \mathbb{Q}_\ell)). \tag{4.379}$$

Comparing (4.379) with (4.371) shows that the numbers λ_j are the eigenvalues of the Frobenius acting on $H^1(\bar{C}, \mathbb{Q}_\ell)$.

The analog of the Riemann Hypothesis for the zeta functions $\zeta_{\mathbb{K}}(s)$ of function fields was stated in 1924 by E. Artin as the property that the zeros lie on the line $\Re(s) = 1/2$, or equivalently that the complex numbers λ_i of (4.372) satisfy

$$|\lambda_j| = \sqrt{q}. \tag{4.380}$$

The statement was proved in 1933 by Hasse in the case of elliptic curves and in 1940 by Weil for all smooth curves C over a finite field. This motivated the formulation of an higher dimensional analog (the Weil conjectures), which in turn greatly stimulated the development of the theory of motives and of a great part of the landscape of modern algebraic geometry. The Weil conjectures were proved by Deligne in 1974.

The Weil proof in the case of curves can be formulated either using étale cohomology, or purely in terms of the Jacobian of the curve, or again (equivalently) in terms of divisors on $C \times C$. In the rest of this section we are going to analyze some aspects of the Weil proof following this last viewpoint.

6.2. Correspondences and divisors in $C \times C$.

A correspondence Z in this setting is given by a divisor on $C \times C$. Correspondences form a ring under composition given by the usual intersection product formula

$$Z_1 \circ Z_2 = (p_{13})_* (p_{12}^* Z_1 \bullet p_{23}^* Z_2), \tag{4.381}$$

with $p_{ij} : C \times C \times C \to C \times C$ the projections.
This ring has an involution given by

$$Z' = \sigma(Z) \tag{4.382}$$

where σ is the involution that exchanges the two copies of C in the product $C \times C$.

We can assign two degrees to a correspondence Z given by the intersection numbers

$$d(Z) = Z \bullet (P \times C) \quad \text{and} \quad d'(Z) = Z \bullet (C \times P), \quad \forall P \in C. \tag{4.383}$$

We refer to them respectively as the degree $d(Z)$ and the codegree $d'(Z)$.

Lemma 4.138. *The degrees* (4.383) *satisfy*

$$d(Z') = d'(Z), \quad \textit{and} \quad d(Z_1 \circ Z_2) = d(Z_1) d(Z_2). \tag{4.384}$$

PROOF. One has $d(Z') = \sigma(Z) \bullet (P \times C) = Z \bullet (C \times P) = d'(Z)$. Moreover, $d(Z_1 \circ Z_2)$ is the degree of the divisor on C defined by the intersection of $Z_1 \circ Z_2$ with $P \times C$. The intersection $p_{12}^* Z_1 \cap p_{23}^* Z_2$ consists of points $(P, Q = Z_2(P), Z_1(Q))$ in $C \times C \times C$, which the pushforward along p_{13} maps to pairs $(P, Z_1(Z_2(P)))$ with multiplicity given by $Z_2 \bullet (P \times C) = d(Z_2)$, so that we get

$$d(Z_1 \circ Z_2) = d(Z_2)\, Z_1 \bullet (Q \times C) = d(Z_1) d(Z_2).$$

□

The correspondences $P \times C$ and $C \times P$ are called *trivial correspondences*. One can consider the abelian group $\mathrm{Div}_{tr}(C \times C)$ generated by these trivial correspondence and take the quotient

$$\mathcal{C}(C) := \mathrm{Div}(C \times C)/\mathrm{Div}_{tr}(C \times C). \tag{4.385}$$

This is still endowed with the ring structure given by composition.
It is always possible to change the degree and codegree of a correspondence Z by adding a multiple of the trivial correspondences $P \times C$ and $C \times P$, so that, for any element in $\mathcal{C}$ we find a representative $Z \in \mathrm{Corr}$ with

$$d(Z) = d'(Z) = 0. \tag{4.386}$$

We also want to consider correspondences up to the linear equivalence relation on divisors in $C \times C$. Namely, the equivalence

$$Z_1 \sim Z_2 \Longleftrightarrow Z_1 - Z_2 = \mathrm{div}(f), \tag{4.387}$$

where $\mathrm{div}(f)$ is a principal divisor on $C \times C$. Thus, one can consider

$$\mathrm{Pic}(C \times C) = \mathrm{Div}(C \times C)/\sim$$

and its quotient $\mathcal{P}(C)$ modulo the classes of the trivial correspondences. This is still a ring under composition.
A correspondence Z is effective if it is given by an effective divisor on $C \times C$, namely if it is a combination $Z = \sum_i n_i Z_i$ of curves $Z_i \subset C \times C$ with coefficients $n_i \geq 0$. We write $Z \geq 0$ to mean its effectiveness.
An effective correspondence $Z \geq 0$ can be viewed as a multivalued map

$$Z : C \to C, \quad P \mapsto Z(P), \tag{4.388}$$

of which the divisor is the graph. (With a slight abuse of notation we will use the same letter Z for the multivalued map and its graph.)
In the case of effective correspondences, one has a product $\star$ given by the composition of multivalued maps

$$Z = Z_1 \star Z_2, \quad Z_1 \star Z_2(P) = Z_1(Z_2(P)), \quad \forall P \in C. \tag{4.389}$$

LEMMA 4.139. *Notice that the composition* (4.389) *agrees with the usual composition* (4.381)*,*

$$Z_1 \star Z_2 = Z_1 \circ Z_2. \tag{4.390}$$

PROOF. The value $Z_2(P) = Z_1(Z_2(P))$ of the multivalued functions is given by the intersection product of their graphs as above. □

In the following, for simplicity, we just write $Z_1 \star Z_2(P) = Z_1(Z_2(P))$ for the product of correspondences, regardless of their being effective. Lemma 4.139 clarifies the meaning of this notation.

DEFINITION 4.140. *The trace of a correspondence Z on $C \times C$ is the expression*

$$\mathrm{Tr}(Z) = d(Z) + d'(Z) - Z \bullet \Delta, \tag{4.391}$$

with Δ the diagonal (identity correspondence) and $\bullet$ the intersection product.

The formula gives $\mathrm{Tr}(Z) = 0$ for principal divisors and for trivial correspondences, hence it descends to the quotient.

6.3. Frobenius correspondences and effective divisors.

We discuss in more detail the step of changing a correspondence $Z = \sum a_n \mathrm{Fr}^n$ of degree g given by a combination of powers of the Frobenius to an effective correspondence through principal divisors.

LEMMA 4.141. *Let C be a smooth projective curve over a finite field and let D be a divisor on C of degree one, which is contained in a given orbit of the Frobenius. In general, it is not possible to modify D to an effective divisor in such a way that the added principal divisor has support on the same orbit.*

PROOF. We exhibit an explicit example, by considering the case of an elliptic curve. Then we have a pair $Z = \{u, \bar{u}\}$ of elements in $S^1 = \{z \in \mathbb{C} \,|\, |z| = 1\}$ such that

$$1 + q - q^{1/2}(u + \bar{u}) = N, \tag{4.392}$$

where N is the number of points of the curve with coordinates in $\mathbb{F}_q$.

Consider in particular the case of the curve C over $\mathbb{F}_4$ of equation

$$y^2 + y = x^3 + \zeta, \tag{4.393}$$

where ζ is a non-trivial cubic root of 1 in $\mathbb{F}_4$. The homogeneous equation of the curve is

$$f(x, y, z) = y^2 z + y z^2 + x^3 + \zeta z^3 = 0. \tag{4.394}$$

One checks that C is non-singular. In fact, $\partial_x f = 0$ gives $x = 0$, $\partial_y f = 0$ gives $z = 0$, and at $P = (0, 1, 0)$ one has $\partial_z f \neq 0$. The point P is the only point at infinity.

Moreover, for $z \neq 0$ no point of the curve is defined over $\mathbb{F}_4$. In fact, one has $y^2 + y \in \mathbb{F}_2$ for all $y \in \mathbb{F}_4$, hence $y^2 + y + \zeta \neq 0$. Thus, since the multiplicative group of $\mathbb{F}_4$ has order 3, the equation $y^2 + y + \zeta = x^3$ has no solution. (Since we work in characteristic two, signs are irrelevant.)

Thus, the number (4.392) of points of the curve is $N = 1$. One has $1 + q - N = 4$, hence $2(u + \bar{u}) = 4$, which implies $u = 1$. This gives for the number of points of the curve over $\mathbb{F}_{4^n}$ the value

$$N(n) = 4^n - 2 \times 2^n + 1. \tag{4.395}$$

For instance, for $n = 2$ and $n = 3$, one gets 9 points of the curve over $\mathbb{F}_{16}$ and 49 points over $\mathbb{F}_{64}$, respectively.

One has to be careful about the normalization factor of $q^{1/2} = 2$ when looking at the action of the Frobenius on the ℓ-adic cohomology. Without this factor this action would have two eigenvalues equal to 1 and would be unipotent, while it cannot be trivial since we can view it as the action of the Frobenius on the torsion points of the curve. With the correct normalization factor one sees that in fact it coincides with multiplication by 2.

This can be seen easily on the 9 points of the curve over $\mathbb{F}_{16}$. One can see that these are the 3-torsion points of the curve. On them the Frobenius action has to agree with the symmetry $X \mapsto -X$, since only the neutral element P is fixed. In fact, modulo three the multiplication by 2 is equal to this symmetry. Similarly one can check that the 49 points over $\mathbb{F}_{64}$ are the 7-torsion of the curve while on them the Frobenius acts by multiplication by 2.

Let Q be any of the 48 points over $\mathbb{F}_{64}$ different from the neutral element P. The orbit of the Frobenius acting on Q is $\{2^n Q\}$ where the meaning of the multiplication by 2 is as above, i.e. $2Q = Q+Q$ where the addition takes place in the elliptic curve. Thus the orbit of the Frobenius here has three elements Q, $2Q$, $4Q$ (one has $8Q = Q$ since Q is a 7-torsion element).

Let D be the divisor of degree one

$$D = \sum_j n_j P_j = 2 \cdot (2Q) - Q, \tag{4.396}$$

which is the difference of $2Q$ with coefficient 2 and Q with coefficient one. Then, by Abel's theorem, the only effective divisor linearly equivalent to D is the point $3Q$ but this point does not belong to the same Frobenius orbit. Thus $D-(3Q)$, which is the principal divisor needed to make D effective, is not of the form $\sum a_n \mathrm{Fr}^n(Q)$. □

The example described above shows that one usually cannot transform a correspondence $Z = \sum a_n \mathrm{Fr}^n$ of positive degree into an effective correspondence by adding a principal correspondence that is also supported on the orbit of the Frobenius action. However, the following result shows that one can achieve effectiveness while remaining in the commutant of the Frobenius.

PROPOSITION 4.142. *A correspondence $Z = \sum a_n \mathrm{Fr}^n$ of degree g on a curve C over $\mathbb{F}_q$ can be made effective by adding a principal correspondence which is defined over $\mathbb{F}_q$ and which commutes with* Fr.

PROOF. The Riemann–Roch theorem ensures that one can make Z effective by adding a principal correspondence, over the field $k(P)$, where k is the common field of definition of the correspondence Z and of the curve (cf. [**278**]) and P is a generic point. A correspondence of the form $Z = \sum a_n \mathrm{Fr}^n$ is in fact defined over $\mathbb{F}_q$, hence the principal correspondence is also defined over $\mathbb{F}_q$. As such it automatically commutes with Fr (cf. [**292**], p. 287). □

6.4. Positivity in the Weil proof.

The main steps in the Weil proof of RH for function fields are

(1) The explicit formula
(2) Positivity

We have talked about explicit formulae in Chapter 2 and in the previous sections of this chapter. We now concentrate on the issue of positivity.

Weil positivity here consists of the following statement

THEOREM 4.143. *Let Z be a correspondence on $C \times C$ modulo the trivial correspondences and up to linear equivalence. Then if Z is nontrivial it satisfies*

$$\mathrm{Tr}(Z \star Z') > 0. \tag{4.397}$$

PROOF. The first step in the proof consists of using the Riemann–Roch formula on C to show that one can achieve effectiveness.

LEMMA 4.144. *There is a representative Z in the class modulo trivial correspondences and linear equivalence that satisfies*

$$Z > 0 \quad \text{and} \quad d(Z) = g. \tag{4.398}$$

PROOF. One first uses the trivial correspondences to adjust the degree to be $d(Z) = g$. This is always possible as the degree is additive and $d(C \times P) = 1$. One then uses the Riemann–Roch formula (4.368) for the curve C. This shows that if D is a divisor on C of degree $\deg(D) = g$ then $\ell(D) > 0$, i.e. there are $f \in \mathbb{K}^*$ such that $\mathrm{div}(f) + D \geq 0$, which means that there are effective representatives in the linear equivalence class of D.
The intersection of $Z \subset C \times C$ with $P \times C$ defines a divisor $Z(P)$ on C with

$$\deg(Z(P)) = d(Z) = g.$$

Thus, the argument above shows that there exists $f_P \in \mathbb{K}^*$ such that $Z(P) + (f_P)$ is effective. This determines an effective divisor $Z + (f)$ on $C \times C$. ☐

We will now continue using the notation Z for the effective divisor on $C \times C$ obtained in Lemma 4.144. Since Z is effective, we can write is as a multivalued function

$$P \mapsto Z(P) = Q_1 + \cdots + Q_g. \tag{4.399}$$

LEMMA 4.145. *The product $Z \star Z'$ is of the form*

$$Z \star Z' = d'(Z)\Delta + Y, \tag{4.400}$$

where Δ is the diagonal in $C \times C$ and Y is the effective correspondence such that $Y(P)$ is the divisor on C given by the sum of points in $\{Q \in C |\, Q = Q_i(P) = Q_j(P),\ i \neq j\}$.

PROOF. One sees from the description in terms of intersection product that it is given by the multivalued function

$$(Z \star Z')(Q) = \sum_{i,j} \sum_{P \in \mathcal{U}_{ij}(Q)} P, \tag{4.401}$$

where

$$\mathcal{U}_{ij}(Q) = \{P \in C |\, Q_i(P) = Q_j(P) = Q\}.$$

One can then separate this out in the contribution of the locus where $Q_i = Q_j$ for $i \neq j$ and the part where $i = j$,

$$(Z \star Z')(Q) = U(Q) + Y(Q)$$

Notice that

$$\#\{P \in C |\, Q = Q_i(P),\ \text{for some } i = 1, \ldots, g\} = d'(Z). \tag{4.402}$$

Thus, for $i = j$ we obtain that the divisor $U(Q) = \sum_i \sum_{P \in \mathcal{U}_{ii}(Q)} P$ is just $d'(Z)\Delta(Q)$, while for $i \neq j$ one obtains the remaining term Y of (4.400). ☐

In the case $g = 1$, the effective correspondence $Z(P) = Q(P)$ is single valued and the divisor $(Z \star Z')(P)$ of (4.401) reduces to the sum of points in $\mathcal{U}(Q) = \{P \in C \,|\, Q(P) = Q\}$. There are $d'(Z)$ such points so one obtains

$$Z \star Z' = d'(Z)\Delta, \quad \text{with} \quad \mathrm{Tr}(Z \star Z') = 2d'(Z) \geq 0, \tag{4.403}$$

since for $g = 1$ one has $\Delta \bullet \Delta = 0$ and $d'(Z) \geq 0$ since Z is effective.

In the case of genus $g > 1$, the Weil proof proceeds as follows. Let κ_C be a choice of an effective canonical divisor for C without multiple points, and let $\{f_1, \ldots, f_g\}$ be a basis of the space $H^0(\kappa_C)$. One then considers the function $C \to M_{g\times g}(\mathbb{F}_q)$ to $g \times g$ matrices

$$P \mapsto M(P), \quad \text{with} \quad M_{ij}(P) = f_i(Q_j(P)). \tag{4.404}$$

and the function $K : C \to \mathbb{F}_q$ given by

$$K(P) = \det(M(P))^2. \tag{4.405}$$

Lemma 4.146. *The function $P \mapsto K(P)$ of (4.405) is a rational function with $(2g-2)d'(Z)$ double poles.*

Proof. The function $K(P)$ is a symmetric function of the $Q_j(P)$, because of squaring the determinant. The composition $P \mapsto (Q_j(P)) \mapsto K(P)$ is then a rational function of $P \in C$.

The poles occur (as double poles) at those points $P \in C$ for which some $Q_i(P)$ is a component of κ_C. The canonical divisor κ_C has degree $2g-2$. This means that there are $(2g-2)d'(Z)$ such double poles. ☐

This yields the following result.

Lemma 4.147. *For $Z \star Z' = d'(Z)\Delta + Y$ as above, the intersection number $Y \bullet \Delta$ satisfies the estimate*

$$Y \bullet \Delta \leq (4g-4)\, d'(Z). \tag{4.406}$$

Proof. By Lemma 4.146, the rational function $K(P)$ of (4.405) has a number of zeros equal to $(4g-4)\, d'(Z)$. On the other hand, $Y \bullet \Delta$ counts the number of times that $Q_i = Q_j$ for $i \neq j$. Since each point P with $Q_i(P) = Q_j(P)$ for $i \neq j$ produces a zero of $K(P)$, one sees that $Y \bullet \Delta$ satisfies the estimate (4.406). ☐

Notice that, for genus $g > 1$ the self intersection of the diagonal is the Euler characteristic

$$\Delta \bullet \Delta = 2 - 2g = \chi(C). \tag{4.407}$$

Moreover, we have

$$d(Z \star Z') = d(Z)d'(Z) = g\, d'(Z) = d'(Z \star Z'). \tag{4.408}$$

Thus using again the decomposition (4.400) and the definition of the trace of a correspondence (4.391), together with (4.407) and (4.408) one obtains

$$\begin{aligned} \mathrm{Tr}(Z \star Z') &= 2g\, d'(Z) + (2g-2)\, d'(Z) - Y \bullet \Delta \\ &\geq (4g-2)\, d'(Z) - (4g-4)\, d'(Z) = 2d'(Z) \geq 0. \end{aligned} \tag{4.409}$$

This completes the proof of Theorem 4.143. ☐

To end the proof of RH for function fields, one applies Theorem 4.143 to a particular type of correspondences, namely those that are of the form

$$Z_{n,m} = m\Delta + n\mathrm{Fr}, \tag{4.410}$$

for $n, m \in \mathbb{Z}$, with Fr the Frobenius correspondence.

Notice that, while the correspondence depends linearly on $n, m \in \mathbb{Z}$, the expression for the trace gives

$$\mathrm{Tr}(Z_{n,m} \star Z'_{n,m}) = 2gm^2 + 2(1+q-N)mn + 2gqn^2, \tag{4.411}$$

where $N = \#C(\mathbb{F}_q)$.

In particular, (4.411) depends quadratically on (n, m). In the process of passing from a correspondence of degree g to an effective correspondence, this quadratic dependence on (n, m) is contained in the multiplicity $d'(Z)$. In fact, one can see this in the following simple example, where one considers a correspondence $Z_{0,m} = m\Delta$.

Corollary 4.148. *On an elliptic curve C consider the correspondence $Z_{0,m} = m\Delta$. The codegree $d'(Z)$ depends quadratically on m.*

Proof. We change $Z_{0,m}$ to make it of degree one, without changing the codegree, by subtracting the trivial correspondence

$$(m-1)\, C \times O,$$

where O is the origin of the elliptic curve. One then has

$$Z = m\,\Delta - (m-1)\, C \times O,$$

with $d(Z) = 1$. Modulo principal divisors, Z is equivalent to the correspondence $P \mapsto Q(P) = mP$, where mP stands for the sum of m copies of P in the abelian group law of the elliptic curve C. This map is m^2 to 1, so that $d(Q') = d'(Z)$ is a quadratic function of m. $\square$

One further remark is that the argument does not depend on the torsion part of the ring of correspondences. Namely, instead of working with the ring $\mathcal{C}$ of classes of correspondences modulo the trivial ones (or $\mathcal{P}$ modulo linear equivalence), which contains zero divisors, one can work with $\mathcal{C}_{\mathbb{Q}} = \mathcal{C} \otimes_{\mathbb{Z}} \mathbb{Q}$.

7. A noncommutative geometry perspective

This final section reflects our ongoing work with C. Consani aimed at building a dictionary between the main notions and steps involved in the Weil proof of RH for function fields and the noncommutative geometry of the adèle class space.

This is work in progress and much of it is still at a fairly speculative level. We collect here some actual results that identify some connections, and some speculations that give some heuristic justification for parts of our proposed dictionary, which we summarize at the end of the section.

Roughly speaking, the idea is that the role that the geometry of the curve C plays in the Weil proof will be taken over by the complement of the idèle class group $C_{\mathbb{K}}$ inside the noncommutative adèle class space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ (more accurately described in terms of the cyclic module given by the cokernel of the reduction map, i.e. by the motive $\mathcal{H}^1(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*, C_{\mathbb{K}})$ defined above. In the function field case, in fact, one can recover the curve C (up to a non-canonical identification of points in the fiber of the map (4.362) from points $C(\bar{\mathbb{F}}_q)$ to places $\Sigma_{\mathbb{K}}$) as the set of classical points of the periodic orbits for the action of $C_{\mathbb{K}}$ on the adèle class space $X_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$. This locus continues to make sense in the number field case and provides a replacement for the curve C itself. Correspondences on $C \times C$, and especially the Frobenius correspondence, are replaced by operators implementing the $C_{\mathbb{K}}$-action. At the time of this writing, some key notions in this picture still need to be properly identified, such as the analog of principal divisors and linear equivalence.

7.1. Distributional trace of a flow.

We present here the computation of the distributional trace of a flow on a smooth manifold, due to [**156**] which in turn can be seen as a generalization of the usual Atiyah–Bott–Lefschetz formula for a diffeomorphism of a smooth compact manifold [**9**]. We first define the distributional trace using the formulation of the functorial properties of distributions ([**156**] VI §1).

Let us first recall the basic notion of generalized section of a vector bundle on a manifold.

Definition 4.149. *Given a vector space E over $\mathbb{R}$, $\dim E = n$, a density $\rho \in |E^*|$ is a map*

$$\rho : \bigwedge^n E \to \mathbb{C} \tag{4.412}$$

such that $\rho(\lambda v) = |\lambda|\, \rho(v)$ for all $\lambda \in \mathbb{R}$ and for all $v \in \bigwedge^n E$.

Given a linear map $L : E \to F$ we let $|L^*| : |F^*| \to |E^*|$ be the corresponding linear map

$$\rho \mapsto \rho \circ \bigwedge^n L\,. \tag{4.413}$$

It depends contravariantly on L. A smooth compactly supported density $\rho \in C_c^\infty(M, |T^*M|)$ on an arbitrary manifold M has a canonical integral

$$\int \rho \in \mathbb{C}\,. \tag{4.414}$$

Definition 4.150. *One defines the space of generalized sections of a vector bundle L on M as*

$$C^{-\infty}(M, L) = \text{ dual of } C_c^\infty(M, L^* \otimes |T^*M|) \tag{4.415}$$

where L^* is the dual bundle.

One has a natural inclusion

$$C^\infty(M, L) \subset C^{-\infty}(M, L) \tag{4.416}$$

given by the pairing

$$\sigma \in C^\infty(M, L)\ ,\ s \in C_c^\infty(M, L^* \otimes |T^*M|) \to \int \langle s, \sigma\rangle \tag{4.417}$$

where $\langle s, \sigma\rangle$ is viewed as a density, $\langle s, \sigma\rangle \in C_c^\infty(M, |T^*M|)$ using the duality pairing of L and L^* (i.e. if $\sigma = u \otimes \rho$ then $\langle s, \sigma\rangle = u(s)\rho$). One has a similar notion of generalized section with compact support, denoted by $C_c^{-\infty}(M, L)$.

The Schwartz kernel theorem is the following statement.

THEOREM 4.151. *Any continuous linear map*

$$T : C_c^\infty(Y) \to C^{-\infty}(X)\,. \tag{4.418}$$

can be written as

$$(T\,\xi)\,(x) = \int k(x,y)\,\xi(y)\,dy \tag{4.419}$$

where $k(x,y)\,dy$ is a generalized section

$$k \in C^{-\infty}(X \times Y,\ \mathrm{pr}_Y^*(|T^*Y|))\,. \tag{4.420}$$

Under suitable transversality conditions the following distribution theoretic trace makes sense:

DEFINITION 4.152. *The distribution theoretic trace of*

$$T : C_c^\infty(X) \to C^{-\infty}(X)$$

is given by

$$\mathrm{Tr}_{\mathrm{distr}}(T) = \int k(x,x)\,dx\,. \tag{4.421}$$

where k is the associated Schwartz kernel.

One obtains ([**156**] VI §1) a generalized section from a submanifold and a symbol as follows.

LEMMA 4.153. *Let L be a vector bundle on the manifold X, $Z \subset X$ a submanifold, and*

$$\sigma \in C^\infty(Z, L \otimes |N_Z|)\,, \tag{4.422}$$

where N_Z is the normal bundle of Z. Then the equality

$$\delta_{Z,\sigma}(\xi) = \int_Z \sigma\,\xi\,, \quad \forall \xi \in C_c^\infty(X, L^* \otimes |T^*X|) \tag{4.423}$$

determines a generalized section $\delta_{Z,\sigma}$ of the vector bundle L on the manifold X.

PROOF. The product $\sigma\,\xi|_Z$ is a density on Z, since it is a section of $|T_Z^*| = |T_X^*| \otimes |N_Z|$. One can thus integrate it over Z. □

LEMMA 4.154. *Let $\varphi : X \to Y$ be smooth and transverse to the submanifold $Z \subset Y$. The pull back $\varphi^*\,\delta_{Z,\sigma}$ of the distribution $\delta_{Z,\sigma}$ makes sense and it is equal to $\delta_{\varphi^{-1}(Z),\,\varphi^*\sigma}$, where*

$$\varphi^*\sigma(x) = |(\varphi_*)^{-1}|\,\sigma(\varphi(x)) \in (\varphi^*\,L)_x \otimes |N_x|\,, \quad \forall x \in \varphi^{-1}(Z)\,. \tag{4.424}$$

Here $|(\varphi_*)^{-1}|$ is the inverse determinant of the isomorphism of normal bundles

$$\varphi_* :\, N_{\varphi^{-1}(Z)} \sim \varphi^*\,N_Z\,. \tag{4.425}$$

PROOF. We assume that φ is transverse to Z, so that for each $x \in X$ with $y = \varphi(x) \in Z$ one has

$$\varphi_*(T_x) + T_{\varphi(x)}(Z) = T_y\,Y\,. \tag{4.426}$$

Let

$$\tau_x = \{X \in T_x\ ,\ \varphi_*(X) \in T_y(Z)\}\,. \tag{4.427}$$

Then φ_* gives a canonical isomorphism of normal bundles

$$\varphi_* : N_x(\varphi^{-1}(Z)) = T_x(X)/\tau_x \sim T_y(Y)/T_y(Z) = N_{\varphi(x)}(Z)\,. \tag{4.428}$$

□

We next consider a manifold M with a vector field $v \in C^\infty(M, T_M)$, and we assume the existence of the corresponding flow

$$F_t = \exp(tv) \quad \forall t \in \mathbb{R}. \tag{4.429}$$

We get a one-parameter group of operators acting on $C^\infty(M)$,

$$(U_t\,\xi)(x) = \xi(F_t(x)) \qquad \forall\,\xi \in C^\infty(M)\,,\ x \in M\,,\ t \in \mathbb{R}\,. \tag{4.430}$$

Since $(F_t)_*\mu v_x = \mu v_x$, the tangent map $(F_t)_*$ induces the Poincaré return map

$$(F_t)_* : T_x/\mathbb{R}\,v_x \to T_x/\mathbb{R}\,v_x = N_x \tag{4.431}$$

THEOREM 4.155. *([**156**] VI §1) Assume that the support of the test function h is contained in the subset where the following transversality assumption holds*

$$1 - (F_t)_* : T_x/\mathbb{R}\,v_x \to T_x/\mathbb{R}\,v_x \quad \textit{is invertible.} \tag{4.432}$$

Then the distributional trace is given by

$$\mathrm{Tr}_{\mathrm{distr}}\left(\int h(t)\,U_t\,dt\right) = \sum_\gamma \int_{I_\gamma} \frac{h(u)}{|1-(F_u)_*|}\,d^*u\,, \tag{4.433}$$

where the sum is over the periodic orbits γ (including the zeros $x \in M$, $v_x = 0$), $I_\gamma \subset \mathbb{R}$ is the isotropy subgroup of any $x \in \gamma$, and d^*u is the unique Haar measure in I_γ such that the covolume of I_γ is equal to 1.

To be more precise, the normalization of d^*u is such that, for the unique Haar measure $d\mu$ of total mass 1 on $\mathbb{R}/I_\gamma$ and for any $f \in C_c^\infty(\mathbb{R})$, it satisfies

$$\int_{\mathbb{R}} f(t)\,dt = \int_{\mathbb{R}/I_\gamma}\left(\int_{I_\gamma} f(u+s)\,d^*u\right)d\mu(s)\,.$$

PROOF. Let us now consider a manifold M with a flow (4.429) and the corresponding map f given by

$$f : M \times \mathbb{R} \to M\ ,\ f(x,t) = F_t(x)\,. \tag{4.434}$$

The graph of f is the submanifold Z of $Y = M \times \mathbb{R} \times M$

$$Z = \{(x,t,y)\,|\,y = F_t(x)\}\,. \tag{4.435}$$

The Schwartz kernel k of the operator

$$(T\,\xi)\,(x) = \xi\,(f(x)) \ \ \forall \xi \in C_c^\infty(M)\,, \tag{4.436}$$

is the distribution $\delta_{Z,\sigma} = \delta(y - f(x))dy$ on Y such that

$$\int \delta(y - f(x))\,\xi(y)\,dy = \xi(f(x)) \ \ \forall \xi \in C_c^\infty(M)\,. \tag{4.437}$$

We apply the above discussion with $X = M \times \mathbb{R}$, and the submanifold $Z \subset Y$. One lets φ be the diagonal map

$$\varphi(x,t) = (x,t,x)\ ,\ \varphi : X \to Y\,. \tag{4.438}$$

The transversality condition $\varphi \pitchfork Z$ holds at $(x,t) \in X$ iff every element of T_xM is of the form

$$(F_t)_* X - X + \mu\, v_x\,, \quad X \in T_xM\,,\ \mu \in \mathbb{R}\,, \tag{4.439}$$

and the transversality condition (4.432) means exactly that (4.439) holds. By definition, the left hand side of the distributional trace formula (4.433) is given by the pushforward

$$q_*\varphi^*(\delta(y - f(x))\,dy)\,, \tag{4.440}$$

where q is the projection $q(x,t) = t \in \mathbb{R}$. The distribution $\varphi^*(\delta(y - f(x))\,dy)$ on X is carried by the $(x,t) \in X = M \times \mathbb{R}$ such that $\varphi(x,t) \in Z$, i.e. such that $F_t(x) = x$. Thus, except for $t = 0$, it is carried by the union of periodic orbits. These correspond either to a zero $v_x = 0$ of the vector field, or a periodic orbit with period $T^\#$.
In the first case, the transversality condition (4.439) becomes

$$1 - (F_t)_* \ \text{ invertible at } x\,. \tag{4.441}$$

One has $F_t(x) = x$, for all $t \in \mathbb{R}$, and the relevant component W of $\varphi^{-1}(Z)$ is $\{x\}\times\mathbb{R}$. The transverse space $N(W)$ is identified with T_x and the map $\varphi_* : N(W) \sim N(Z)$ is given by

$$\varphi_* = 1 - (F_t)_*\,. \tag{4.442}$$

Thus the symbol of $\varphi^*(\delta(y - f(x))\,dy)$ is the scalar function $|1 - (F_t)_*|^{-1}$. The generalized section $q_*\,\varphi^*(\delta(y - F_t(x))\,dy)$ is the function $t \mapsto |1 - (F_t)_*|^{-1}$. We can then write the contribution of the zeros of the flow as

$$\sum_{\text{zeros}} \int \frac{h(t)}{|1 - (F_t)_*|}\, dt \tag{4.443}$$

where h is a test function vanishing at 0.
We then look at the second case i.e. the contribution of a periodic orbit. The corresponding part of $\varphi^{-1}(Z)$ is of the form

$$\varphi^{-1}(Z) = V \times \Gamma \subset M \times \mathbb{R}\,, \tag{4.444}$$

where Γ is a discrete cocompact subgroup of $\mathbb{R}$, while $V \subset M$ is a one-dimensional compact submanifold of M. The computation of the contribution of the component $V \times \{T\}$ of $\varphi^{-1}(Z)$ to (4.440) gives

$$T^\# \frac{1}{|1 - (F_T)_*|}\, h(T)\,, \tag{4.445}$$

where $T^\#$ is the length of the primitive orbit or equivalently the covolume of Γ in $\mathbb{R}$ for the Haar measure $|dt|$. We can thus write the contributions of the periodic orbits as

$$\sum_{\gamma_p}\sum_{\Gamma} \mathrm{Covol}(\Gamma)\, \frac{1}{|1 - (F_T)_*|}\, h(T)\,, \tag{4.446}$$

where the test function h vanishes at 0. When collecting the various contributions (4.443) and (4.446) one obtains the required formula (4.433). $\Box$

7.2. The periodic orbits of the action of $C_{\mathbb{K}}$ on $X_{\mathbb{K}}$.

We explain here the results of [**71**], Section VI and Appendix III. They generalize the analogous formula for flows on manifolds of [**156**] explained above in §7.1.
We let $\mathbb{K}$ be a global field and $X_{\mathbb{K}} = \mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ the adèle class space. The action of the idèle class group $C_{\mathbb{K}}$ on $X_{\mathbb{K}}$ is given by the the analog

$$m : X_{\mathbb{K}} \times C_{\mathbb{K}} \to X_{\mathbb{K}} \ , \ m(x,u) = u\,x\,. \tag{4.447}$$

of the map f of (4.434). The graph of m is the subspace

$$Z = \{(x,u,y)\,|\,y = u\,x\}\,. \tag{4.448}$$

of $Y = X_{\mathbb{K}} \times C_{\mathbb{K}} \times X_{\mathbb{K}}$. The diagonal map φ is given by

$$\varphi(x,u) = (x,u,x) \ , \ \varphi : X_{\mathbb{K}} \times C_{\mathbb{K}} \to Y\,. \tag{4.449}$$

The periodic orbits are obtained from the pullback

$$\varphi^{-1}(Z) = \{(x,u) \in X_{\mathbb{K}} \times C_{\mathbb{K}}\,|\,u\,x = x\} \tag{4.450}$$

PROPOSITION 4.156. *Let $(x,u) \in \varphi^{-1}(Z)$, with $u \neq 1$. There exists a place $v \in \Sigma_{\mathbb{K}}$ such that*

$$x \in X_{\mathbb{K},v} = \{x \in X_{\mathbb{K}}\,|\,x_v = 0\} \tag{4.451}$$

The isotropy subgroup of any $x \in X_{\mathbb{K},v}$ contains the cocompact subgroup

$$\mathbb{K}_v^* \subset C_{\mathbb{K}}\,, \quad \mathbb{K}_v^* = \{(k_w)\,|\,k_w = 1 \ \forall w \neq v\} \tag{4.452}$$

PROOF. Let $(x,u) \in X_{\mathbb{K}} \times C_{\mathbb{K}}$ with $ux = x$. This equality still holds for suitable lifts $(\tilde{x},\tilde{u}) \in \mathbb{A}_{\mathbb{K}} \times \mathbb{A}_{\mathbb{K}}^*$. Since $u \neq 1$ by hypothesis, one has $\tilde{u} \neq 1$ and there exists a place $v \in \Sigma_{\mathbb{K}}$ such that $\tilde{u}_v \neq 1$. Then the equality $\tilde{u}\tilde{x} = \tilde{x}$ shows that $(\tilde{u}_v - 1)\tilde{x}_v = 0$ and $\tilde{x}_v = 0$.
One has $ux = x$ for any $u \in \mathbb{K}_v^*$ and any $x \in X_{\mathbb{K},v}$. It remains to show that $\mathbb{K}_v^*$ is cocompact in $C_{\mathbb{K}}$. It is a closed subgroup since the group homomorphism $\mathbb{K}_v^* \to C_{\mathbb{K}}$ is proper as can be seen since it preserves the modules of these modulated groups. In fact the module of an idèle (k_w) is the product $\prod |k_w|$. □

REMARK 4.157. This shows that the action of $C_{\mathbb{K}}$ on $X_{\mathbb{K}}$ has the peculiar property that the non-trivial isotropy groups always contain a cocompact subgroup. In fact Proposition 4.156 shows that the $I_v = \mathbb{K}_v^*$ are the *minimal* non-trivial isotropy subgroups. They are cocompact and pairwise disjoint.

The transverse space to $X_{\mathbb{K},v} \subset X_{\mathbb{K}}$ is identified with the transverse space of the closed $\mathbb{K}^*$-invariant subset of adèles $\mathbb{A}_{\mathbb{K},v} \subset \mathbb{A}_{\mathbb{K}}$,

$$\mathbb{A}_{\mathbb{K},v} = \{a = (a_w)_{w\in\Sigma_{\mathbb{K}}}\,|\,a_v = 0\}. \tag{4.453}$$

and hence with the local field $\mathbb{K}_v$. The action of the isotropy group $I_v = \mathbb{K}_v^*$ is simply the action of $\mathbb{K}_v^*$ on $\mathbb{K}_v$ by multiplication and the corresponding term in the trace formula (4.433) is given by

$$\int_{\mathbb{K}_v^*} \frac{h(u^{-1})}{|1-u|}\, d^*u \tag{4.454}$$

where the u^{-1} in the argument of h comes from the notation for the action of $C_{\mathbb{K}}$ on $X_{\mathbb{K}}$,

$$\vartheta_{\mathrm{a}}(u)\xi(x) = \xi(u^{-1}x), \tag{4.455}$$

which differs from (4.430) by the presence of the inverse.

7.3. Frobenius (scaling) correspondence and the trace formula.

We begin by reformulating the geometric side of the trace formula given in Theorem 4.130 in more intersection theoretic language, so as to be able to compare the terms in the trace formula with the setup of §6.2 above.

To the scaling action $\vartheta_{\mathrm{a}}(\gamma)(\xi)(x) = \xi(\gamma^{-1}x)$ one associated the graph Z_g given by the pairs $(x, g^{-1}x)$. These should be considered as points in the product $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^* \times \mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ of two copies of the adèle class space. Thus, the analog in our context of the correspondences $Z = \sum_n a_n \mathrm{Fr}^n$ on $C \times C$ considered in §6.3 above is given by elements of the form

$$Z(f) = \int_{C_{\mathbb{K}}} f(g)\, Z_g\, d^*g, \tag{4.456}$$

for some $f \in \mathbf{S}(C_{\mathbb{K}})$.

With this interpretation of correspondences, we can then make sense of the terms in the trace formula in the following way.

Definition 4.158. *For a correspondence of the form* (4.456) *we define degree and codegree by the following prescription:*

$$d(Z(f)) := \hat{f}(1) = \int_{C_{\mathbb{K}}} f(u)\, |u|\, d^*u, \tag{4.457}$$

$$d'(Z(f)) := d(Z(f^\sharp)) = \int_{C_{\mathbb{K}}} f(u)\, d^*u = \hat{f}(0). \tag{4.458}$$

Here the Fourier transform $\hat{f}$ is as in (4.295), with the trivial character $\chi = 1$. Notice that, with this definition of degree and codegree we find

$$d(Z_g) = |g|, \qquad \text{and} \qquad d'(Z_g) = 1. \tag{4.459}$$

Thus, the term $\hat{f}(1) + \hat{f}(0)$ in the trace formula of Theorem 4.130 matches the term $d(Z) + d'(Z)$ in Weil's formula for the trace of a correspondence as in (4.391). The term

$$\int'_{(\mathbb{K}_v^*, \alpha_v)} \frac{f(u^{-1})}{|1-u|}\, d^*u \tag{4.460}$$

of (4.346) in turn can be seen as the remaining term $Z \bullet \Delta$ in (4.391). In fact, as we saw in §7.2 the formula (4.460) describes, using distributions, the local contributions to the trace of the intersections between the graph $Z(f)$ and the diagonal Δ.

When we separate out, as in Lemma 4.131, the contribution $\log|a|\, h(1)$ and we rewrite the trace formula as in Theorem 4.130. This corresponds to separating the intersection $Z \bullet \Delta$ into a term that is proportional to the self-intersection $\Delta \bullet \Delta$ and a remaining term, where the intersection is transverse.

To see this, we notice that the term $\log|a|$, for $a = (a_v)$ a differental idèle, is of the form (4.344). Indeed one sees that, in the function field case the term

$$-\log|a| = -\log q^{2-2g} = (2g-2)\log q = -\Delta \bullet \Delta \, \log q$$

is proportional to the self-intersection of the diagonal, which brings us to consider the value $\log|a| = -\log D$ with the discriminant of a number field as the analog in characteristic zero of the self-intersection of the diagonal.

In these intersection theoretic terms we can reformulate the positivity condition equivalent to the Riemann Hypothesis in the following way.

PROPOSITION 4.159. *The following two conditions are equivalent:*

- *All L-functions with Grössencharakter on $\mathbb{K}$ satisfy the Riemann Hypothesis.*
- *The estimate*

$$Z(f) \bullet_{\text{trans}} Z(f) \leq 2d(Z(f))d'(Z(f)) - \Delta \bullet \Delta\, f \star f^\sharp(1) \tag{4.461}$$

holds for all $f \in \mathbf{S}(C_{\mathbb{K}})$.

PROOF. As in the Weil proof one separates the terms

$$Z \star Z' = d'(Z)\Delta + Y,$$

where Y has transverse intersection with the diagonal. Here we can write an identity

$$\mathrm{Tr}(\underline{\vartheta}_{\mathrm{m}}(f \star f^\sharp)|H^1) =: Z(f) \bullet Z(f) = \Delta \bullet \Delta\, f \star f^\sharp(1) + Z(f) \bullet_{\text{trans}} Z(f), \tag{4.462}$$

where the remaning term $Z(f) \bullet_{\text{trans}} Z(f)$, which represents the transverse intersection, is given by the local contributions given by the principal values over $(\mathbb{K}_v^*, e_{\mathbb{K}_v})$ in the formula (4.345) for $\mathrm{Tr}(\underline{\vartheta}_{\mathrm{m}}(f \star f^\sharp)|H^1)$.

The formula (4.345) for $\mathrm{Tr}(\underline{\vartheta}_{\mathrm{m}}(f \star f^\sharp)|H^1)$ gives a term of the form

$$-\log|a|\, f \star f^\sharp(1),$$

with

$$f \star f^\sharp(1) = \int_{C_{\mathbb{K}}} |f(g)|^2\, |g|\, d^*g.$$

We rewrite this term as $-\Delta \bullet \Delta\, f \star f^\sharp(1)$ according to our interpretation of $\log|a|$ as the self-intersection of the diagonal. This matches the term $(2g-2)d'(Z)$ in the estimate for $\mathrm{Tr}(Z \star Z')$ in the Weil proof.

The first two terms in the formula (4.345) for $\mathrm{Tr}(\underline{\vartheta}_{\mathrm{m}}(f \star f^\sharp)|H^1)$ are of the form

$$\widehat{f \star f^\sharp}(0) + \widehat{f \star f^\sharp}(1) = 2\hat{f}(0)\hat{f}(1) = 2d'(Z(f))d(Z(f)). \tag{4.463}$$

This matches the term $2gd'(Z) = 2d(Z)d'(Z)$ in the expression for $\mathrm{Tr}(Z \star Z')$ in the Weil proof.

With this notation understood, we see that the positivity

$$\mathrm{Tr}(\underline{\vartheta}_{\mathrm{m}}(f \star f^\sharp)|H^1) \geq 0$$

indeed corresponds to the estimate (4.461). ∎

7.4. The Fubini theorem and trivial correspondences.

The first important step in the Weil proof, as we have seen in §6.2 above, is the freedom to modify the degree by adding multiples of the trivial correspondences.

In our setting, what plays the role of the trivial correspondences, in view of the vanishing result of Proposition 4.135 should be given by the correspondences $Z(f)$ associated to elements $f \in \mathcal{V} \subset \mathbf{S}(C_{\mathbb{K}})$, in the range of the reduction map from adèles. Thus, we need to show that it is possible to arbitrarily modify the degree $d(Z(f)) = \hat{f}(1)$ of a correspondence by adding to f an element in $\mathcal{V}$.

This is possible (as we show in Lemma 4.160 below) only thanks to a failure of Fubini's theorem. In fact, at first sight it would appear that, if we consider an element $\xi \in \mathcal{S}(\mathbb{A}_{\mathbb{K}})_0$, since this satisfies the vanishing condition

$$\int_{\mathbb{A}_{\mathbb{K}}} \xi(x)\, dx = 0,$$

one should also obtain that, for the function on $C_{\mathbb{K}}$ defined by $f(x) = \sum_{k \in \mathbb{K}^*} \xi(kx)$, the integral

$$\hat{f}(1) = \int_{C_{\mathbb{K}}} f(g)|g|d^*g \tag{4.464}$$

should vanish. Indeed, for local fields (but not in the global case) the relation between the additive and multiplicative Haar measures is of the form $dg = |g|d^*g$. To see more clearly what happens, let us just restrict to the case where $\mathbb{K} = \mathbb{Q}$ and take the function $\xi(x)$ of the form

$$\xi = 1_{\hat{\mathbb{Z}}} \otimes \eta,$$

with $\eta \in \mathcal{S}(\mathbb{R})_0$. We then have $C_{\mathbb{Q}} = \hat{\mathbb{Z}}^* \times \mathbb{R}_+^*$ and the function f is of the form

$$f(u, \lambda) = \sum_{n \in \mathbb{Z}, n \neq 0} \eta(n\lambda), \quad \forall \lambda \in \mathbb{R}_+^*\,,\ u \in \hat{\mathbb{Z}}^*\,. \tag{4.465}$$

We can thus write (4.464) in this case as

$$\hat{f}(1) = \int_{\hat{\mathbb{Z}}^* \times \mathbb{R}_+^*} f(u, \lambda) du\, d\lambda = \int_{\mathbb{R}} \sum_{n \in \mathbb{N}} \eta(n\lambda)\ d\lambda \tag{4.466}$$

Moreover since $\eta \in \mathcal{S}(\mathbb{R})_0$ we have for all n,

$$\int_{\mathbb{R}} \eta(n\lambda) d\lambda = 0\,. \tag{4.467}$$

It is however *not* necessarily the case that we can apply Fubini's theorem and write

$$\int_{\mathbb{R}} \sum_{n \in \mathbb{N}} \eta(n\lambda)\ d\lambda = \sum_n \int_{\mathbb{R}} \eta(n\lambda) d\lambda = 0 \tag{4.468}$$

since as soon as $\eta \neq 0$ one has

$$\sum_{n=1}^{\infty} \int_{\mathbb{R}} |\eta(n\lambda)| d\lambda = (\int_{\mathbb{R}} |\eta(\lambda)| d\lambda)\ \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

so that Fubini's theorem does not apply and one cannot interchange the integral and the sum in (4.468). Thus, one can in general have $\hat{f}(1) \neq 0$, even though $\sum_n \int_{\mathbb{R}} \eta(n\lambda) d\lambda = 0$. In fact, we have the following result.

LEMMA 4.160. *Given $f \in \mathbf{S}(C_{\mathbb{K}})$, it is possible to change arbitrarily the value of the degree $d(Z(f)) = \hat{f}(1)$ by adding elements of $\mathcal{V}$.*

PROOF. It suffices to exhibit an element $f \in \mathcal{V}$ such that $\hat{f}(1) \neq 0$, as then by linearity one can obtain the result. We only treat the case $\mathbb{K} = \mathbb{Q}$. We take $\eta \in \mathcal{S}(\mathbb{R})_0$ given by

$$\eta(x) = \pi x^2(\pi x^2 - \frac{3}{2})\, e^{-\pi x^2}$$

We get by Proposition 2.24 of Chapter 2 that $\mathfrak{E}(1_{\hat{\mathbb{Z}}} \otimes \eta)$ has the Riemann ξ function of Chapter 2, Proposition 2.24 as its Fourier transform. This function does not vanish at $\pm\frac{i}{2}$ thus one gets that $\hat{f}(1) \neq 0$. $\square$

7.5. The curve inside the adèle class space.

In the case of a function field $\mathbb{K} = \mathbb{F}_q(C)$, one can identify a locus inside the adèle class space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$, described as a union of orbits of the $C_{\mathbb{K}}$ action, which is (non-canonically) isomorphic to the algebraic points $C(\bar{\mathbb{F}}_q)$. This locus can be defined in the same way for all global fields and provides a replacement for the curve C inside the adèle class space in the case of a number field. As we show below in §7.5.1, it can also be described as the set of classical points of the periodic orbits $X_{\mathbb{K},v}$ of (4.451) parameterized by the places of $\mathbb{K}$.

DEFINITION 4.161. *Let $\mathbb{K}$ be a global field. For a place $v \in \Sigma_{\mathbb{K}}$ consider the adèle*

$$a^{(v)} = (a_w^{(v)}), \qquad \text{with} \quad a_w^{(v)} = \begin{cases} 1 & w \neq v \\ 0 & w = v. \end{cases} \tag{4.469}$$

The set of periodic classical *points of the adèle class space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ is defined as the union of orbits*

$$\Xi_{\mathbb{K}} := \bigcup_{v \in \Sigma_{\mathbb{K}}} C_{\mathbb{K}} a^{(v)}. \tag{4.470}$$

As before, we denote by $C_{\mathbb{K},1} \subset C_{\mathbb{K}}$ the kernel of the norm $|\cdot| : C_{\mathbb{K}} \to \mathbb{R}_+^*$. In the function field case, we have the following (non-canonical) identification.

PROPOSITION 4.162. *For $\mathbb{K} = \mathbb{F}_q(C)$ a function field, the orbits of Frobenius on $C(\bar{\mathbb{F}}_q)$ give an equivariant identification*

$$\Xi_{\mathbb{K}}/C_{\mathbb{K},1} \simeq C(\bar{\mathbb{F}}_q), \tag{4.471}$$

between $\Xi_{\mathbb{K}}/C_{\mathbb{K},1}$ with the action of $q^{\mathbb{Z}}$ and $C(\bar{\mathbb{F}}_q)$ with the action of the group of integer powers of the Frobenius.

PROOF. At each place $v \in \Sigma_{\mathbb{K}}$ the quotient group of the range N of the norm $|\cdot| : C_{\mathbb{K}} \to \mathbb{R}_+^*$ by the range N_v of $|\cdot| : \mathbb{K}_v \to \mathbb{R}_+^*$ is the finite cyclic group

$$N/N_v = q^{\mathbb{Z}}/q^{n_v\mathbb{Z}} \simeq \mathbb{Z}/n_v\mathbb{Z}, \tag{4.472}$$

where n_v is the degree of the place $v \in \Sigma_{\mathbb{K}}$, cf. §6.1 above. The degree n_v is the same as the cardinality of the orbit of the Frobenius acting on the fiber of the map (4.362) from algebraic points in $C(\bar{\mathbb{F}}_q)$ to places in $\Sigma_{\mathbb{K}}$. Thus, one can construct in this way an equivariant embedding

$$C(\bar{\mathbb{F}}_q) \hookrightarrow (\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*)/C_{\mathbb{K},1} \tag{4.473}$$

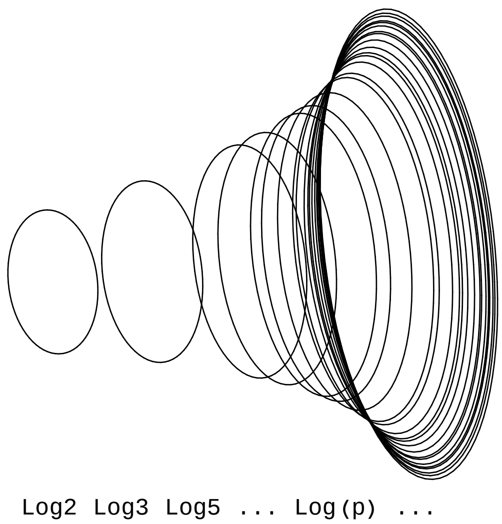


FIGURE 3. The classical periodic points $\Xi_{\mathbb{Q}}/C_{\mathbb{Q},1}$ of the adèle class space.

obtained, after choosing a point in each orbit, by mapping the orbit of the integer powers of the Frobenius in $C(\bar{\mathbb{F}}_q)$ over a place v to the orbit of $C_{\mathbb{K}}/C_{\mathbb{K},1} \sim q^{\mathbb{Z}}$ on the adèle $a^{(v)}$. □

Modulo the problem created by the fact that the identification above is non-canonical and relies upon the choice of a point in each orbit, it is then possible to think of the locus $\Xi_{\mathbb{K}}$, in the number field case, as a replacement for $C(\bar{\mathbb{F}}_q)$ inside the adèle class space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$. It is very important to view the space $\Xi_{\mathbb{K}}$ as embedded in the adèle class space $X_{\mathbb{K}}$ with the induced structure. In fact, it is the embedding $\Xi_{\mathbb{K}} \subset X_{\mathbb{K}}$ that provides the geometric set on which the Lefschetz trace formula of Theorem 4.130. This involves not only the space $\Xi_{\mathbb{K}}$ *per se* but it crucially involves the way this sits inside $X_{\mathbb{K}}$. If we ignore the structure induced by $X_{\mathbb{K}}$, then the space $\Xi_{\mathbb{K}}$ would just be a disjoint union of orbits without any interesting global structure.

The reduction to eigenspaces of Grössencharakters corresponds to induced bundles over the quotient $(\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*)/C_{\mathbb{K},1}$. Thus, we can in fact consider the quotient space $\Xi_{\mathbb{K}}/C_{\mathbb{K},1}$ endowed with the action of the quotient group $C_{\mathbb{K}}/C_{\mathbb{K},1}$.

In the case of $\mathbb{K} = \mathbb{Q}$, the quotient $\Xi_{\mathbb{Q}}/C_{\mathbb{Q},1}$ appears as a union of periodic orbits of period $\log p$ under the action of $C_{\mathbb{Q}}/C_{\mathbb{Q},1} \sim \mathbb{R}$, as in Figure 3.

7.5.1. *The curve and the periodic orbits.*

Each of the periodic orbits $X_{\mathbb{K},v}$ of Proposition 4.156 is still a noncommutative space. By construction the space $\Xi_{\mathbb{K}}$ of Definition 4.161 is a union over the set $\Sigma_{\mathbb{K}}$ of places of $\mathbb{K}$ of the subsets

$$\Xi_{\mathbb{K},v} = C_{\mathbb{K}} a^{(v)} \subset X_{\mathbb{K},v}\,. \tag{4.474}$$

It is thus important to understand how one can deduce $\Xi_{\mathbb{K},v}$ in a natural manner from the noncommutative space $X_{\mathbb{K},v}$. The space $X_{\mathbb{K},v}$ comes naturally with a measure, and hence an associated time evolution which we now compute. Indeed, by construction one has

$$X_{\mathbb{K},v} = \mathbb{A}_{\mathbb{K},v}/\mathbb{K}^* \tag{4.475}$$

where $\mathbb{A}_{\mathbb{K},v}$ is the closed $\mathbb{K}^*$-invariant subset of $\mathbb{A}_{\mathbb{K}}$ defined in (4.453) by the condition $a_v = 0$. This defines an "hyperplane" in $\mathbb{A}_{\mathbb{K}}$ and the additive Haar measure gives a natural measure.

PROPOSITION 4.163. *The modular automorphism group of the additive Haar measure on the crossed product* $\mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K},v}$ *is given by the time evolution*

$$\sigma_t^v(f)(k,x) = |k|_v^{it}\, f(k,x), \quad \forall t \in \mathbb{R}, \quad \forall f \in C^*(\mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K},v})\,. \tag{4.476}$$

PROOF. We identify elements of $C_c(\mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K},v})$ with functions $f(g,a)$ of elements $g \in \mathbb{K}^*$ and $a \in \mathbb{A}_{\mathbb{K},v}$. The product is simply of the form

$$f_1 * f_2(g,a) = \sum_r f_1(g\,r^{-1}, ga) f_2(r,a).$$

The additive Haar measure da on $\mathbb{A}_{\mathbb{K},v}$ satisfies the scaling property

$$d(ka) = |k|_v^{-1} da\,, \quad \forall k \in \mathbb{K}^*\,, \tag{4.477}$$

since the product measure $da \times da_v$ on $\mathbb{A}_{\mathbb{K}} = \mathbb{A}_{\mathbb{K},v} \times \mathbb{K}_v$ is invariant under the scaling by $k \in \mathbb{K}^*$, while the additive Haar measure da_v on $\mathbb{K}_v$ gets multiplied by $|k|_v$: $d(ka_v) = |k|_v da_v$. We then check the KMS_1 condition, for the weight φ associated to the additive Haar measure, as

$$\varphi(f_1 * f_2) = \sum_r \int_{\mathbb{A}_{\mathbb{K},v}} f_1(r^{-1}, ra) f_2(r,a)\, da$$

$$= \sum_r \int_{\mathbb{A}_{\mathbb{K},v}} f_2(k^{-1}, kb) f_1(k,b)\, |k|_v^{-1}\, db = \varphi(f_2 * \sigma_i(f_1))\,,$$

using the change of variables $k = r^{-1}$, $a = kb$ and $da = |k|_v^{-1}\, db$. □

In order to deal with states rather than weights, we replace the groupoid $\mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K},v}$ by the Morita equivalent groupoid obtained by restricting to the open subset

$$\mathbb{A}^{(1)}_{\mathbb{K},v} \subset \mathbb{A}_{\mathbb{K},v} = \prod_{w \neq v} \mathbb{K}^{(1)}_w \tag{4.478}$$

where for the local field $\mathbb{K}_w$ we let $\mathbb{K}^{(1)}_w$ be the *interior* of $\{x \in \mathbb{K}_w\,;\, |x| \leq 1\}$. We let $\mathbb{K}^* \ltimes \mathbb{A}^{(1)}_{\mathbb{K},v}$ be the restricted groupoid

$$\mathcal{G}(v) = \mathbb{K}^* \ltimes \mathbb{A}^{(1)}_{\mathbb{K},v} = \{(g,a) \in \mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K},v} \,|\, a \text{ and } ga \in \mathbb{A}^{(1)}_{\mathbb{K},v}\}\,. \tag{4.479}$$

We still denote by φ the restriction of the above weight to $C^*(\mathbb{K}^* \ltimes \mathbb{A}^{(1)}_{\mathbb{K},v})$. It is a positive functional given by

$$\varphi(f) = \int_{\mathbb{A}^{(1)}_{\mathbb{K},v}} f(1,a)\, da. \tag{4.480}$$

Proposition 4.163 shows that its modular automorphism group σ_t^φ is given by (4.476).

LEMMA 4.164. *Let $\mathbb{K}$ be a global field. Let $W \subset \mathbb{A}_{\mathbb{K}}$ be a neighborhood of $0 \in \mathbb{A}_{\mathbb{K}}$. Then for $x \in \mathbb{A}_{\mathbb{K}}$ one has $\mathbb{K}^* x \cap W \neq \emptyset$, unless $x \in \mathbb{A}_{\mathbb{K}}^*$ is an idèle. For $x \in \mathbb{A}_{\mathbb{K}}^*$, the orbit $\mathbb{K}^* x$ is discrete in $\mathbb{A}_{\mathbb{K}}$.*

PROOF. One can assume that W is of the form

$$W = \{a = (a_w) |\ |a_w| < \varepsilon \ \forall w \in S \text{ and } |a_w| \leq 1 \ \forall w \notin S\},$$

for S a finite set of places and for some $\varepsilon > 0$. Multiplying by a suitable idèle one can in fact assume that $S = \emptyset$, so that we have

$$W = \{a = (a_w) |\ |a_w| \leq 1 \ \forall w \in \Sigma_{\mathbb{K}}\}.$$

One has $|x_v| \leq 1$ except on a finite set $F \subset \Sigma_{\mathbb{K}}$ of places. Moreover, if x is not an idèle, one can also assume that

$$\prod_{v \in F} |x_v| < \delta$$

for any fixed δ. Thus, $-\log |x_v|$ is as large as one wants and there exists $k \in \mathbb{K}^*$ such that $kx \in W$. This is clear in the function field case because of the Riemann–Roch formula (4.368). In the case of $\mathbb{Q}$ one can first multiply x by an integer to get $|x_v| \leq 1$ for all finite places, then since this does not alter the product of all $|x_v|$ one gets $|x_\infty| < 1$ and $x \in W$. The case of more general number fields is analogous. In the case of idèles, one can assume that $x = 1$ and then the second statement follows from the discreteness of $\mathbb{K}$ in $\mathbb{A}_{\mathbb{K}}$. ☐

We now recall that, for an étale groupoid like $\mathcal{G}(v)$, every unit $y \in \mathcal{G}(v)^{(0)}$ defines, by (3.43) of Chapter 3, i.e.

$$(\pi_y(f)\xi)(\gamma) = \sum_{\gamma_1 \gamma_2 = \gamma} f(\gamma_1)\xi(\gamma_2), \tag{4.481}$$

a representation π_y by left convolution of the algebra of $\mathcal{G}(v)$ in the Hilbert space $\mathcal{H}_y = \ell^2(\mathcal{G}(v)_y)$, where $\mathcal{G}(v)_y$ denotes the set of elements of the groupoid $\mathcal{G}(v)$ with source y. By construction, the unitary equivalence class of the representation π_y is unaffected when one replaces y by an equivalent $z \in \mathcal{G}(v)^{(0)}$, i.e. one assumes that there exists $\gamma \in \mathcal{G}(v)$ with range and source y and z. Thus, we can think of the label y of π_y as living in the quotient space $X_{\mathbb{K},v}$ of equivalence classes of elements of $\mathcal{G}(v)^{(0)}$.

The relation between $\Xi_{\mathbb{K},v}$ and $X_{\mathbb{K},v}$ is then the following.

THEOREM 4.165. *Let $y \in X_{\mathbb{K},v}$, the representation π_y is a positive energy representation if and only if $y \in \Xi_{\mathbb{K},v}$.*

PROOF. Let first $y \in \mathcal{G}(v)^{(0)} \cap \Xi_{\mathbb{K},v}$. Thus one has $y \in \mathbb{A}^{(1)}_{\mathbb{K},v}$, $y_w \neq 0$ for all w and $|y_w| = 1$ for all $w \notin S$ where S is a finite set of places. We can identify $\mathcal{G}(v)_y$ with the set of $k \in \mathbb{K}^*$ such that $ky \in \mathbb{A}^{(1)}_{\mathbb{K},v}$. We extend y to the adèle $\tilde{y} = y \times 1$ whose component at the place v is equal to $1 \in \mathbb{K}_v$. Then $\tilde{y}$ is an idèle. Thus, by Lemma 4.164, the number of elements of the orbit $\mathbb{K}^*\tilde{y}$ in a given compact subset of $\mathbb{A}_{\mathbb{K}}$ is finite. It follows that $\log |k|_v$ is lower bounded on $\mathcal{G}(v)_y$. Indeed, otherwise there would exist a sequence $k_n \in \mathbb{K}^* \cap \mathcal{G}(v)_y$ such that $|k_n|_v \to 0$. Then $k_n \tilde{y} \in \mathbb{A}^{(1)}_{\mathbb{K}}$ for all

n large enough and this contradicts the discreteness of $\mathbb{K}^*\tilde{y}$. In the representation π_y the time evolution σ_t is implemented by the Hamiltonian

$$(H_y\,\xi)(k,y) = \log|k|_v\ \xi(k,y). \tag{4.482}$$

Namely, we have

$$\pi_y(\sigma_t(f)) = e^{itH_y}\pi_y(f)e^{-itH_y}, \quad \forall f \in C_c(\mathcal{G}(v))\,. \tag{4.483}$$

Thus, since $\log|k|_v$ is lower bounded on $\mathcal{G}(v)_y$, we get that the representation π_y is a positive energy representation.

Let then $y \in \mathcal{G}(v)^{(0)} \smallsetminus \Xi_{\mathbb{K},v}$. We shall show that $\log|k|_v$ is not lower bounded on $\mathcal{G}(v)_y$, and thus that π_y is not a positive energy representation. We consider as above the adèle $\tilde{y} = y \times 1$ whose component at the place v is equal to $1 \in \mathbb{K}_v$. Assume that $\log|k|_v$ is lower bounded on $\mathcal{G}(v)_y$. Then there exists $\epsilon > 0$ such that, for $k \in \mathbb{K}^*$,

$$k\,y \in \mathbb{A}^{(1)}_{\mathbb{K},v} \Rightarrow |k|_v \geq \epsilon\,.$$

This shows that the neighborhood of $0 \in \mathbb{A}_{\mathbb{K}}$ defined as

$$W = \{a \in \mathbb{A}_{\mathbb{K}}\,;\, |a_v| < \epsilon\,,\ a_w \in \mathbb{K}_w^{(1)}\,,\ \forall w \neq v\}$$

does not intersect $\mathbb{K}^*\tilde{y}$. Thus by Lemma 4.164 we get that $\tilde{y}$ is an idèle and $y \in \Xi_{\mathbb{K},v}$. □

The specific example of the Bost-Connes system combined with Theorem 4.165 shows that we can refine our recipe for taking "classical points" of a noncommutative space. The latter recipe only provides a notion of classical points that can be thought of, by analogy with the positive characteristic case, as points defined over the mysterious "field with one element" $\mathbb{F}_1$ (see e.g. [**218**]). To obtain instead a viable notion of the points defined over the maximal unramified extension $\bar{\mathbb{F}}_1$, one performs the following sequence of operations.

$$X \overset{\text{Dual System}}{\longrightarrow} \hat{X} \overset{\text{Periodic Orbits}}{\longrightarrow} \cup\,\hat{X}_v \overset{\text{Classical Points}}{\longrightarrow} \cup\,\Xi_v, \tag{4.484}$$

which make sense in the framework of endomotives of §2. Note in particular that the dual system $\hat{X}$ is of type II and as such does not have a non-trivial time evolution.

Thus it is only by restricting to the periodic orbits that one passes to noncommutative spaces of type III for which the cooling operation is non-trivial. In the analogy with geometry in non-zero characteristic, the set of points $X(\bar{\mathbb{F}}_q)$ over $\bar{\mathbb{F}}_q$ of a variety X is indeed obtained as the union of the periodic orbits of the Frobenius.

REMARK 4.166. Theorem 4.165 does not give the classification of KMS_β states for the quantum statistical system $(C^*(\mathbb{K}^* \ltimes \mathbb{A}^{(1)}_{\mathbb{K},v}), \sigma_t)$. It only allows one to construct extremal KMS_β states of the form

$$\psi_{\beta,\,y}(a) = \mathrm{Tr}(e^{-\beta H_y}\,\pi_y(a))/\mathrm{Tr}(e^{-\beta H_y})\,, \quad \forall a \in C^*(\mathcal{G}(v))\,. \tag{4.485}$$

7.5.2. *The valuation systems for* $\mathbb{K} = \mathbb{Q}$.

We analyzed in Chapter 3 several examples of quantum statistical systems for which the structure of KMS_β states stabilized at low enough temperature. In the present case of the valuation systems $(C^*(\mathcal{G}(v)), \sigma_t)$ a very interesting new phenomenon occurs already when $\mathbb{K} = \mathbb{Q}$ and v is a finite place. Indeed, while the elements of $\Xi_{\mathbb{K}, v}$ label extremal KMS_β states $\psi_{\beta, y}$ for $\beta > 1$, given by (4.485), it is no longer true that these states converge to a unique vacuum state $\psi_{\infty, y}$ for $\beta \to \infty$ and there is a finite degeneracy of the vacuum state. We now explain in detail what happens in this case $\mathbb{K} = \mathbb{Q}$.

Definition 4.167. *Let* $\mathbb{A}^*_{\mathbb{Q},p} \subset \mathbb{A}^{(1)}_{\mathbb{Q},p}$ *be the subspace*

$$\mathbb{A}^*_{\mathbb{Q},p} = \{x \in \mathbb{A}_{\mathbb{Q},p} |\ |x_w| = 1\ \forall w \neq p, \infty \quad \text{and } p^{-1} \leq |x_\infty| < 1\}. \tag{4.486}$$

As above $\mathcal{G}(p)$ denotes the reduction (4.479) of the groupoid $\mathbb{Q}^* \ltimes \mathbb{A}_{\mathbb{Q},p}$ by the open subset $\mathbb{A}^{(1)}_{\mathbb{Q},p} \subset \mathbb{A}_{\mathbb{Q},p}$, namely

$$\mathcal{G}(p) = \{(k, x) \in \mathbb{Q}^* \ltimes \mathbb{A}_{\mathbb{Q},p} \,|\, x \in \mathbb{A}^{(1)}_{\mathbb{Q},p},\ kx \in \mathbb{A}^{(1)}_{\mathbb{Q},p}\}. \tag{4.487}$$

Notice that the set $\mathbb{A}^{(1)}_{\mathbb{Q},p}$ meets all the equivalence classes in $\mathbb{A}_{\mathbb{Q},p}$ by the action of $\mathbb{Q}^*$. In fact, given $x \in \mathbb{A}_{\mathbb{Q},p}$, one can find a representative y with $y \sim x$ in $\mathbb{A}_{\mathbb{Q},p}/\mathbb{Q}^*$, such that $y \in \hat{\mathbb{Z}} \times \mathbb{R}$. Upon multiplying y by a suitable power of p, one can make y_∞ as small as required, and in particular one can obtain in this way a representative in $\mathbb{A}^{(1)}_{\mathbb{Q},p}$. Let us assume that $|y_w| = 1$ for all finite places $w \neq p$ and that $y_\infty > 0$. Then there exists a unique $n \in \mathbb{N} \cup \{0\}$ such that $p^n\, y \in \mathbb{A}^*_{\mathbb{Q},p}$.
Given a prime p we define the function $f_p(\lambda, \beta)$ for $\lambda \in (1, p]$ and $\beta > 1$ by

$$f_p(\lambda, \beta) = \sum c_k\, p^{-k\beta} \tag{4.488}$$

where the $c_k \in \{0, \dots p-1\}$ are the digits of the expansion of λ in base p. There is an ambiguous case where all digits c_k are equal to 0 for $k > m$ while $c_m > 0$, since the same number

$$\lambda = \sum c_k\, p^{-k} \tag{4.489}$$

is obtained using the same c_j for $j < m$, $c_m - 1$ instead of c_m and $c_j = p - 1$ for $j > m$. In that case, for $\beta > 1$, (4.488) gives two different values and we choose the value coming from the second representation of λ, i.e. the lower of the two.

Lemma 4.168. *The coefficients* c_k *of the expansion of* λ *in base* p *are given by*

$$c_k = \lceil \lambda p^k - 1 \rceil - p \lceil \lambda p^{k-1} - 1 \rceil, \tag{4.490}$$

where $\lceil x \rceil = \inf_{n \in \mathbb{Z}} \{n \geq x\}$ *denotes the ceiling function. One has*

$$f_p(\lambda, \beta) = \sum (\lceil \lambda p^k - 1 \rceil - p \lceil \lambda p^{k-1} - 1 \rceil)\, p^{-k\beta}. \tag{4.491}$$

Proof. Indeed, it is enough to show that

$$\lceil \lambda p^k - 1 \rceil = \sum_0^k c_j\, p^{k-j}. \tag{4.492}$$

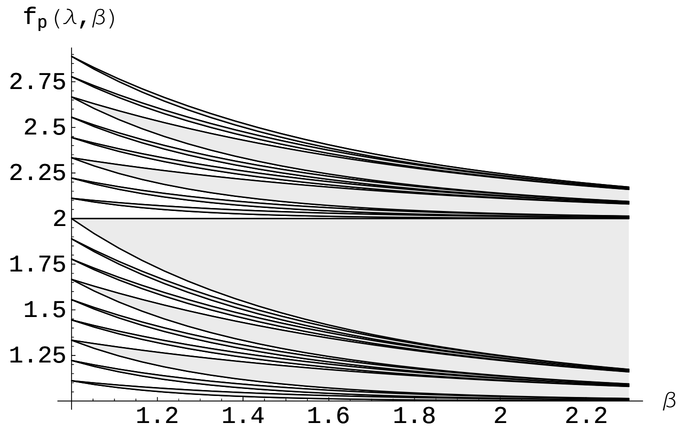


FIGURE 4. Graphs of the functions $f_p(\lambda,\beta)$ as functions of β for $p=3$, $\lambda=n/27$. The gray regions are the gaps in the range of f_p.

Using (4.489), one has

$$\lambda p^k = \sum_0^k c_j\, p^{k-j} + \epsilon\,, \quad \epsilon = \sum_1^\infty c_{k+n}\, p^{-n}\,,$$

where $\epsilon > 0$ since we exclude the case $c_j = 0$ for all $j > k$. One has $\epsilon \le 1$ by construction, thus $\lceil \epsilon - 1 \rceil = 0$ and one gets (4.492) since $\sum_0^k c_j\, p^{k-j}$ is an integer. Then (4.491) follows. ∎

Note that, for $\beta > 1$, the function $f_p(\lambda,\beta)$ is discontinuous (cf. Figures 4 and 5) at any point (λ,β) where the expansion of λ in base p is ambiguous, i.e. $\lambda \in \mathbb{N}\, p^{-k}$. Moreover for $\beta = 1$ one gets

$$f_p(\lambda,1) = \lambda\,, \quad \forall \lambda \in (1,p\,]\,. \tag{4.493}$$

We then obtain the following result.

THEOREM 4.169. *Let $(C^*(\mathcal{G}(p)),\sigma_t^p)$ be the C^*-dynamical system associated to the groupoid (4.487) with the time evolution (4.476). Then the following properties hold.*

(1) *For any $y \in \mathbb{A}^*_{\mathbb{Q},p}$ the corresponding representation π_y has positive energy.*
(2) *Let H_y denote the Hamiltonian implementing the time evolution in the representation π_y, for $y \in \mathbb{A}^*_{\mathbb{Q},p}$ with $y_\infty = \lambda^{-1}$ and $\lambda \in (1,p\,]$. Then the partition function is given by*

$$Z_p(\lambda,\beta) = \mathrm{Tr}(e^{-\beta H_y}) = 2\,\frac{1-p^{-\beta}}{1-p^{1-\beta}}\, f_p(\lambda,\beta)\,. \tag{4.494}$$

(3) *The functionals*

$$\psi_{\beta,y}(a) = \mathrm{Tr}(e^{-\beta H_y}\,\pi_y(a))\,, \quad \forall a \in C^*(\mathcal{G}(p)) \tag{4.495}$$

satisfy the KMS_β condition for σ_t^p and depend weakly continuously on the parameter $y \in \mathbb{A}^*_{\mathbb{Q},p}$.

PROOF. (1) This follows from Theorem 4.165. For $y \in \mathbb{A}^*_{\mathbb{Q},p}$ one has

$$r \in \mathbb{Q}^*,\ ry \in \mathbb{A}^{(1)}_{\mathbb{Q},p} \Longrightarrow r = p^{-k} m, \tag{4.496}$$

for some $k \geq 0$ and some integer m prime to p and such that $|ry_\infty| < 1$. This implies

$$|m| < p^{k+1} \tag{4.497}$$

and one finds

$$|r|_p = p^k \geq 1 \quad \text{and} \quad \log |r|_p \geq 0, \tag{4.498}$$

In fact, the argument above shows that the spectrum of the Hamiltonian H_y implementing the time evolution σ_t^p in the representation π_y is given by

$$\mathrm{Spec}(H_y) = \{k \log p\}_{k \in \mathbb{N} \cup \{0\}}, \tag{4.499}$$

hence π_y is a positive energy representation.
(2) We begin with the special case of $y_\infty = p^{-1}$. Then $\lambda = p$ and $f_p(\lambda, \beta) = \frac{p-1}{1-p^{-\beta}}$ since all digits of $\lambda = p$ are equal to $p - 1$. We want to show that the partition function is given by

$$\mathrm{Tr}(e^{-\beta H_y}) = 2\,\frac{p-1}{1-p^{1-\beta}}. \tag{4.500}$$

The multiplicity of an eigenvalue $k \log p$ of H_y is the number of integers $m \neq 0 \in \mathbb{Z}$ that are prime to p and such that $p^{-k}\,|m|\,y_\infty < 1$. Since we are assuming that $y_\infty = p^{-1}$, this gives $|m| < p^{k+1}$. Thus, the multiplicity is just $2\,(p^{k+1} - p^k)$. The factor 2 comes from the sign of the integer m. The factor $(p^{k+1} - p^k)$ corresponds to subtracting from the number p^{k+1} of positive integers $m \leq p^{k+1}$ the number p^k of those that are multiples of p.

We now pass to the general case. For $x > 0$, $\lceil x - 1 \rceil$ is the cardinality of $(0, x) \cap \mathbb{N}$. The same argument used above shows that the multiplicity of the eigenvalue $k \log p$ is given by the counting

$$2\left(\lceil \lambda p^k - 1 \rceil - \lceil \lambda p^{k-1} - 1 \rceil\right).$$

Thus

$$\mathrm{Tr}(e^{-\beta H_y}) = 2 \sum_{k=0}^{\infty} \left(\lceil \lambda p^k - 1 \rceil - \lceil \lambda p^{k-1} - 1 \rceil\right) p^{-k\beta}. \tag{4.501}$$

One has the following equalities of convergent series:

$$\sum_{k=0}^{\infty} \left(\lceil \lambda p^k - 1 \rceil - \lceil \lambda p^{k-1} - 1 \rceil\right) p^{-k\beta} = \sum_{k=0}^{\infty} \lceil \lambda p^k - 1 \rceil\,(p^{-k\beta} - p^{-(k+1)\beta})$$

so that,

$$\mathrm{Tr}(e^{-\beta H_y}) = 2\,(1 - p^{-\beta}) \sum_{k=0}^{\infty} \lceil \lambda p^k - 1 \rceil\, p^{-k\beta}. \tag{4.502}$$

Similarly

$$\sum_{k=0}^{\infty} \left(\lceil \lambda p^k - 1 \rceil - p\,\lceil \lambda p^{k-1} - 1 \rceil\right) p^{-k\beta} = \sum_{k=0}^{\infty} \lceil \lambda p^k - 1 \rceil\,(p^{-k\beta} - p\,p^{-(k+1)\beta})$$

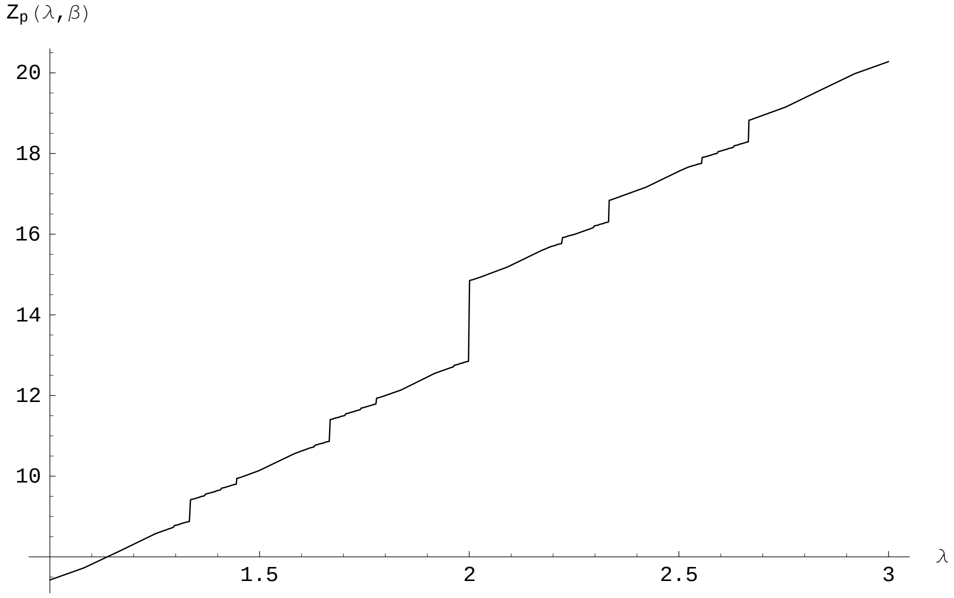


FIGURE 5. Graph of the function $Z_p(\lambda,\beta)$ as a function of λ for $p=3$, $\beta=1.2$.

which gives using (4.491)

$$f_p(\lambda,\beta) = (1-p^{1-\beta})\sum_{k=0}^{\infty}\lceil \lambda p^k - 1\rceil\, p^{-k\beta}\,. \tag{4.503}$$

Combining (4.502) with (4.503) gives (4.494).
(3) It follows from (4.483) and the finiteness of the partition function (4.494) that the functionals (4.495) fulfill the KMS_β condition. In terms of functions on the groupoid $\mathcal{G}(p)$ one has

$$\psi_{\beta,\,y}(f) = \sum f(1, n\, p^{-k}\, y)\, p^{-k\beta}\,, \quad \forall f \in C_c(\mathcal{G}(p)) \tag{4.504}$$

where the sum is absolutely convergent. Each of the terms in the sum gives a weakly continuous linear form. Thus one obtains the required continuity. □

REMARK 4.170. The partition function $Z_p(\lambda,\beta)$ is a discontinuous function of the parameter λ and this might seem to contradict the third statement of Theorem 4.169. It would if the algebra $C^*(\mathcal{G}(p))$ were unital since, in that case, the partition function is given by evaluation on the unit and weak continuity implies that it is continuous. In our case $C^*(\mathcal{G}(p))$ is *not* unital, and the partition function is expressed as a supremum of the form

$$Z_p(\lambda,\beta) = \sup\{\psi_{\beta,\,y}(a^*a) | a \in C^*(\mathcal{G}(p))\,,\ ||a|| \le 1\}\,.$$

In particular it shows that $Z_p(\lambda,\beta)$ is lower semicontinuous as a function of λ.

The precise qualitative properties of the partition functions $Z_p(\lambda,\beta)$ are described by the following result

PROPOSITION 4.171. *As a function of* $\lambda \in (1,p]$ *the partition function* $Z_p(\lambda,\beta)$ *satisfies for* $\beta > 1$*:*

(1) Z_p *is strictly increasing.*
(2) Z_p *is continuous on the left, and lower semicontinuous.*
(3) Z_p *is discontinuous at any point of the form* $\lambda = mp^{-k}$ *with a jump of* $2p^{-k\beta}$ *(for* m *prime to* p*).*
(4) *The measure* $\frac{\partial Z_p}{\partial \lambda}$ *is the sum of the Dirac masses at the points* $\lambda = mp^{-k}$*,* m *prime to* p*, with coefficients* $2p^{-k\beta}$*.*
(5) *The closure of the range of* Z_p *is a Cantor set.*

PROOF. (1) This follows from (4.502) which expresses Z_p as an absolutely convergent sum of multiples of the functions $\lceil \lambda p^k - 1 \rceil$. The latter are non-decreasing and jump by 1 at $\lambda \in \mathbb{N}\, p^{-k} \cap (1,p]$. The density of the union of these finite sets for $k \geq 0$ shows that Z_p is strictly increasing.
(2) This follows as above from (4.502) and the semi-continuity properties of the ceiling function.
(3) Let $\lambda = m\, p^{-k}$ with m prime to p. Then for any $j \geq k$ one gets a jump of $2(1-p^{-\beta})p^{-j\beta}$ coming from (4.502) so that their sum gives

$$2(1-p^{-\beta}) \sum_{j=k}^{\infty} p^{-j\beta} = 2p^{-k\beta}.$$

(4) This follows as above from (4.502) and from (3) which computes the discontinuity at the jumps.
(5) Recall that when writing elements of an interval in base p one gets a map from the Cantor set to the interval. This map is surjective but fails to be injective due to the identifications coming from $\sum_0^\infty (p-1)p^{-m} = p$. The connectedness of the interval is recovered from these identifications. In our case the coefficients c_k of the expansion in base p of elements of $(1,p]$ are such that $c_0 \in \{1,\ldots,p-1\}$ while $c_k \in \{0,\ldots,p-1\}$ for $k > 0$. This is a Cantor set K in the product topology of $K = \{1,\ldots,p-1\} \times \prod_{\mathbb{N}} \{0,\ldots,p-1\}$. As shown in Figure 5, the discontinuities of the function $Z_p(\lambda,\beta)$ as a function of λ replace the connected topology of $(1,p]$ by the totally disconnected topology of K. ∎

REMARK 4.172. One can use (4.501) to define $Z_p(\lambda,\beta)$ for any $\lambda > 0$, as

$$Z_p(\lambda,\beta) = 2\sum_{-\infty}^{\infty} \Big(\lceil \lambda p^k - 1 \rceil - \lceil \lambda p^{k-1} - 1 \rceil \Big)\, p^{-k\beta}. \tag{4.505}$$

This makes sense for $\Re(\beta) > 1$ since $\lceil \lambda p^k - 1 \rceil = 0$ for $k \leq -\frac{\log \lambda}{\log p}$. The extended function (4.505) satisfies

$$Z_p(p\lambda,\beta) = p^\beta Z_p(\lambda,\beta),$$

which suggests replacing $Z_p(\lambda,\beta)$ with

$$\zeta_p(\lambda,\beta) = \lambda^{-\beta} Z_p(\lambda,\beta) \tag{4.506}$$

so that

$$\zeta_p(p\lambda,\beta) = \zeta_p(\lambda,\beta). \tag{4.507}$$

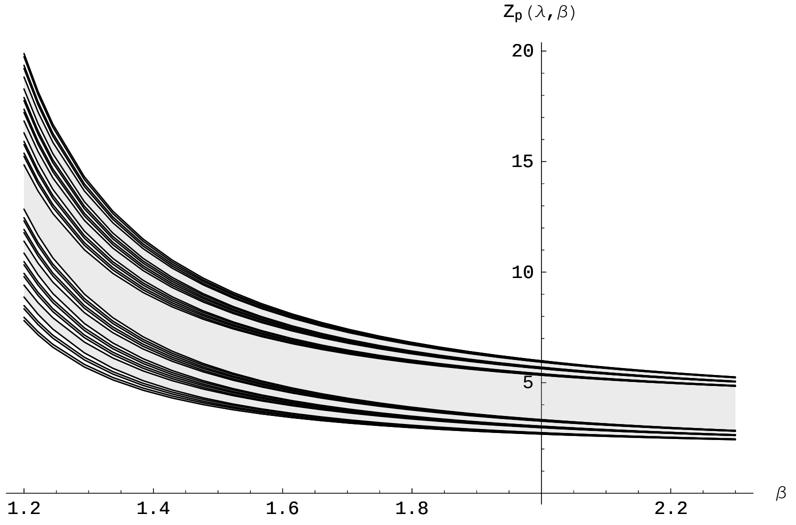


FIGURE 6. Graphs of the functions $Z_p(\lambda, \beta)$ as functions of β for $p = 3$, $\lambda = n/27$. The gray regions are the gaps in the range. All these functions have a pole at $\beta = 1$.

This replacement $Z_p \mapsto \zeta_p$ corresponds to the shift in the Hamiltonian H_y by

$$H_y \mapsto H_y - \log |y_\infty|.$$

We can now refine Theorem 4.165 and consider the zero temperature KMS state of the system $(C^*(\mathcal{G}(p)), \sigma_t^p)$ corresponding to the positive energy representation π_y for $y \in \Xi_{\mathbb{Q},p}$.

PROPOSITION 4.173. *As $\beta \to \infty$ the vacuum states (zero temperature KMS states) of the system $(C^*(\mathcal{G}(p)), \sigma_t^p)$ with Hamiltonian H_y have a degeneracy of $2\lceil \lambda - 1 \rceil$, where $y_\infty = \lambda^{-1}$. There is a preferred choice of a vacuum state given by the evaluation at $y \in \mathbb{A}^*_{\mathbb{Q},p}$.*

PROOF. When we look at the orbit of $y \in \mathbb{A}^*_{\mathbb{Q},p}$, i.e. at the intersection $\mathbb{Q}^* y \cap \mathbb{A}^{(1)}_{\mathbb{Q},p}$ and label its elements by pairs (k, m) as above, we find that all elements with $k = 0$ give a ground state. This degeneracy of the vacuum reflects the fact that the limit of the partition function as the temperature goes to 0 is not in general equal to 1. For instance, for $y_\infty = p^{-1}$, one finds

$$\lim_{\beta \to \infty} \mathrm{Tr}(e^{-\beta H_y}) = \lim_{\beta \to \infty} 2 \frac{p - 1}{1 - p^{1-\beta}} = 2(p - 1).$$

More generally, one finds similarly the limit

$$\lim_{\beta \to \infty} \mathrm{Tr}(e^{-\beta H_y}) = 2\lceil \lambda - 1 \rceil.$$

Among the $2\lceil \lambda - 1 \rceil$ vacuum states, the state given by evaluation at $y \in \mathbb{A}^*_{\mathbb{Q},p}$ is singled out, since $my \notin \mathbb{A}^*_{\mathbb{Q},p}$ for $m \neq 1$. It is then natural to consider, for each finite place $p \in \Sigma_{\mathbb{Q}}$, the section

$$s_p(x) = \mathbb{Q}^* x \cap \mathbb{A}^*_{\mathbb{Q},p}, \quad \forall x \in C_{\mathbb{Q}} a^{(p)} \subset \mathbb{A}_{\mathbb{Q}}/\mathbb{Q}^* \tag{4.508}$$

of the projection from $\mathbb{A}_{\mathbb{Q}}$ to the orbit $C_{\mathbb{Q}} a^{(p)}$. $\square$

Notice that s_p is discontinuous at the boundary of the domain $\mathbb{A}^*_{\mathbb{Q},p}$. Indeed when y_∞ crosses the value p^{-1} the class in $C_{\mathbb{Q}}a^{(p)}$ varies continuously but the representative in $\mathbb{A}^*_{\mathbb{Q},p}$ jumps discontinuously so that its Archimedean component remains in the interval $[p^{-1}, 1)$. This suggests to consider a cyclic covering of $\Xi_{\mathbb{Q}}$, which we now discuss in §7.5.3.

7.5.3. *The cyclic covering $\tilde{\Xi}_{\mathbb{Q}}$ of $\Xi_{\mathbb{Q}}$.*

By construction $\Xi_{\mathbb{K}}$ is a subspace of the adèle class space $X_{\mathbb{K}}$. We shall now show, in the case $\mathbb{K} = \mathbb{Q}$, that it admits a natural lift $\tilde{\Xi}_{\mathbb{Q}}$ to a subspace of $\mathbb{A}_{\mathbb{Q}}$ which reduces the ambiguity group $\mathbb{Q}^*$ to a cyclic group. One thus obtains a natural cyclic covering $\tilde{\Xi}_{\mathbb{Q}} \subset \mathbb{A}_{\mathbb{Q}}$ of $\Xi_{\mathbb{Q}}$. We already saw above, in Proposition 4.173, that it is natural to choose representatives for the elements of the orbit $C_{\mathbb{Q}}a^{(p)}$, for a finite prime p, in the subset of adèles given by

$$\tilde{\Xi}_{\mathbb{Q},p} := \{y \in \mathbb{A}_{\mathbb{Q}} |\; y_p = 0 \text{ and } |y_\ell| = 1 \text{ for } \ell \neq p, \infty \text{ and } y_\infty > 0\}. \tag{4.509}$$

We extend this definition at ∞ by

$$\tilde{\Xi}_{\mathbb{Q},\infty} := \{y \in \mathbb{A}_{\mathbb{Q}} |\; |y_w| = 1 \; \forall w \neq \infty \text{ and } y_\infty = 0\}. \tag{4.510}$$

Definition 4.174. *The locus $\tilde{\Xi}_{\mathbb{Q}} \subset \mathbb{A}_{\mathbb{Q}}$ is defined as*

$$\tilde{\Xi}_{\mathbb{Q}} = \bigcup_{v \in \Sigma_{\mathbb{Q}}} \tilde{\Xi}_{\mathbb{Q},v} \subset \mathbb{A}_{\mathbb{Q}} \tag{4.511}$$

We then have the following simple fact.

Proposition 4.175. *Let π be the projection from $\tilde{\Xi}_{\mathbb{Q}}$ to $\Xi_{\mathbb{Q}}$, with $\pi(x)$ the class of x modulo the action of $\mathbb{Q}^*$.*

1. *The map $\pi : \tilde{\Xi}_{\mathbb{Q}} \to \Xi_{\mathbb{Q}}$ is surjective.*
2. *Two elements in $\tilde{\Xi}_{\mathbb{Q},v}$ have the same image in $C_{\mathbb{Q}}a^{(v)}$ iff they are on the same orbit of the transformation*

$$Tx = px\,, \quad \forall x \in \tilde{\Xi}_{\mathbb{Q},p}\,, \qquad Tx = -x\,, \quad \forall x \in \tilde{\Xi}_{\mathbb{Q},\infty} \tag{4.512}$$

Proof. The first statement follows by lifting $C_{\mathbb{Q}}$ inside $\mathbb{A}^*_{\mathbb{Q}}$ as the subgroup $\hat{\mathbb{Z}}^* \times \mathbb{R}^*_+$. Then any element of $C_{\mathbb{Q}}a^{(v)}$ has a representative in

$$(\hat{\mathbb{Z}}^* \times \mathbb{R}^*_+)a^{(v)}.$$

The proof of the second statement is straightforward, since, for a finite prime p, the subgroup $p^{\mathbb{Z}} \subset \mathbb{Q}^*$ is the group of elements of $\mathbb{Q}^*$ which leave $\tilde{\Xi}_{\mathbb{Q},p}$ globally invariant. □

7.5.4. *The global Morita equivalence.*

We obtained the valuation systems $(C^*(\mathcal{G}(v)), \sigma_t)$ by reducing the groupoids $\mathbb{K}^* \ltimes \mathbb{A}_{\mathbb{K},v}$, associated to the periodic orbits, to the subgroupoids

$$\mathcal{G}(v) = \mathbb{K}^* \ltimes \mathbb{A}^{(1)}_{\mathbb{K},v}.$$

It is important to be able to perform this reduction globally, by reducing the groupoid $\mathcal{G}_{\mathbb{K}}$ to a suitable open set. The set $\mathbb{A}^{(1)}_{\mathbb{K}}$ of (4.513) that we use to reduce the groupoid $\mathcal{G}_{\mathbb{K}}$ will only capture part of the subspace $C_{\mathbb{K}}$, but since our main focus is on the geometry of the complement of this subspace (the cokernel of the reduction map), this will not be a problem.

Definition 4.176. *Consider the open neighborhood of* $0 \in \mathbb{A}_{\mathbb{K}}$ *defined by*

$$\mathbb{A}_{\mathbb{K}}^{(1)} = \prod_{w \in \Sigma_{\mathbb{K}}} \mathbb{K}_w^{(1)} \subset \mathbb{A}_{\mathbb{K}} \tag{4.513}$$

where for any place we let $\mathbb{K}_w^{(1)}$ *be the* interior *of* $\{x \in \mathbb{K}_w \,;\, |x| \leq 1\}$. *Let* $\mathcal{G}_{\mathbb{K}}^{(1)}$ *denote the reduction of the groupoid* $\mathcal{G}_{\mathbb{K}}$ *by the open subset* $\mathbb{A}_{\mathbb{K}}^{(1)} \subset \mathbb{A}_{\mathbb{K}}$ *of the units and* $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$ *denote the corresponding (smooth) groupoid algebra.*

The algebra $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$ is a subalgebra of $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ where one simply extends the function $f(k, x)$ by zero outside of the open subgroupoid $\mathcal{G}_{\mathbb{K}}^{(1)} \subset \mathcal{G}_{\mathbb{K}}$. With this convention, the convolution product of $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$ is simply given by the convolution product of $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ of the form

$$(f_1 \star f_2)(k, x) = \sum_{h \in \mathbb{K}^*} f_1(kh^{-1}, hx) f_2(h, x).$$

We see from Lemma 4.164 above that the only effect of the reduction to $\mathcal{G}_{\mathbb{K}}^{(1)}$ is to remove from the noncommutative space $\mathbb{A}_{\mathbb{K}}/\mathbb{K}^*$ all the elements of $C_{\mathbb{K}}$ whose class modulo $\mathbb{K}^*$ does not intersect $\mathcal{G}_{\mathbb{K}}^{(1)}$ (i.e. in particular those whose norm is greater than or equal to one).

The inclusion $\mathbb{A}_{\mathbb{K},v}^{(1)} \subset \mathbb{A}_{\mathbb{K}}^{(1)}$ is $\mathbb{K}^*$-equivariant and proper, and yields an algebra homomorphism

$$\rho_v \;:\; \mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)}) \to \mathcal{S}(\mathcal{G}_{\mathbb{K},v}^{(1)}) \tag{4.514}$$

which plays the role of the restriction map to the periodic orbit $X_{\mathbb{K},v}$. The modular automorphisms σ_t^v of the valuation systems extend to the global algebra.

Lemma 4.177. *Let* $\mathbb{K}$ *be a global field and* $v \in \Sigma_{\mathbb{K}}$ *a place. The map*

$$d_v(k, x) = \log |k|_v \in \mathbb{R} \tag{4.515}$$

defines a homomorphism of the groupoid $\mathcal{G}_{\mathbb{K}}^{(1)}$ *to the additive group* $\mathbb{R}$ *and the time evolution*

$$\sigma_t^v(f)(k, x) = |k|_v^{it}\, f(k, x), \quad \forall t \in \mathbb{R}, \;\; \forall f \in \mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)}) \tag{4.516}$$

generates a 1-parameter group of automorphisms of the algebra $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$.

The following result shows that the nontrivial part of the dynamics σ_t^v concentrates on the algebra $\mathcal{S}(\mathcal{G}(v))$ with $\mathcal{G}(v)$ as in (4.479).

Proposition 4.178. *The morphism* ρ_v *of* (4.514) *restricts to a* σ_t^v*-equivariant morphism* $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)}) \to \mathcal{S}(\mathcal{G}(v))$. *Moreover, the restriction of the one-parameter group* σ_t^v *to the kernel of* ρ_v *is inner.*

Proof. For the first statement note that the proper inclusion $\mathbb{A}_{\mathbb{K},v} \subset \mathbb{A}_{\mathbb{K}}$ restricts to a proper inclusion $\mathbb{A}_{\mathbb{K},v}^{(1)} \subset \mathbb{A}_{\mathbb{K}}^{(1)}$. For the second statement, notice that the formula

$$h_v(x) = \log |x|_v, \quad \forall x \in \mathbb{A}_{\mathbb{K}}^{(1)}, \tag{4.517}$$

defines the multipliers e^{ith_v} of the kernel of ρ_v. Indeed e^{ith_v} is a bounded continuous function on $\mathbb{A}_{\mathbb{K}}^{(1)} \smallsetminus \mathbb{A}_{\mathbb{K},v}^{(1)}$.

We can then check that the 1-cocycle d_v is the coboundary of h_v. In fact, we have

$$h_v(kx) - h_v(x) = d_v(k,x), \quad \forall (k,x) \in \mathcal{G}_{\mathbb{K}}^{(1)} \smallsetminus \mathcal{G}(v). \tag{4.518}$$

□

We then have the following symmetries for the algebra $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$.

PROPOSITION 4.179. *Let $\mathcal{J}^+$ denote the semigroup of idèles $j \in \mathbb{A}_{\mathbb{K}}^*$ such that $j\mathbb{A}_{\mathbb{K}}^{(1)} \subset \mathbb{A}_{\mathbb{K}}^{(1)}$. The semigroup $\mathcal{J}^+$ acts on the algebra $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$ by endomorphisms obtained as restrictions of the automorphisms of $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ of the form*

$$\vartheta_{\rm a}(j)(f)(k,x) = f(k, j^{-1}x), \quad \forall (k,x) \in \mathcal{G}_{\mathbb{K}}\,, \quad j \in \mathcal{J}^+. \tag{4.519}$$

Let $\mathbb{K} = \mathbb{Q}$ and $C_{\mathbb{Q}}^+ \subset C_{\mathbb{Q}}$ be the semigroup $C_{\mathbb{Q}}^+ = \{g \in C_{\mathbb{Q}} | \, |g| < 1\}$. The semigroup $C_{\mathbb{Q}}^+$ acts on $\mathcal{S}(\mathcal{G}_{\mathbb{Q}}^{(1)})$ by the endomorphisms

$$F(g) = \vartheta_{\rm a}(\bar g), \tag{4.520}$$

with $\bar g$ the natural lift of $g \in C_{\mathbb{Q}}^+$ to $\hat{\mathbb{Z}}^* \times \mathbb{R}_+^*$.

PROOF. By construction $\vartheta_{\rm a}(j)$ is an automorphism of $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$. For a function f with support B in the open set $\mathcal{G}_{\mathbb{K}}^{(1)}$ the support of the function $\vartheta_{\rm a}(j)(f)$ is $jB = \{(k,jx) | (k,x) \in B\} \subset \mathcal{G}_{\mathbb{K}}^{(1)}$, so that $\vartheta_{\rm a}(j)(f)$ still has support in $\mathcal{G}_{\mathbb{K}}^{(1)}$.
For $\mathbb{K} = \mathbb{Q}$ let $\bar g \in \hat{\mathbb{Z}}^* \times \mathbb{R}_+^*$ be the natural lift of an element $g \in C_{\mathbb{Q}}^+$. Then the Archimedean component $\bar g_\infty$ is of absolute value less than 1, so that $\bar g \in \mathcal{J}^+$. The action of $\vartheta_{\rm a}(\bar g)$ by endomorphisms of $\mathcal{S}(\mathcal{G}_{\mathbb{Q}}^{(1)})$ induces a corresponding action of $C_{\mathbb{Q}}^+$. □

REMARK 4.180. For m a positive integer, consider the element $g = (1, m^{-1}) \in C_{\mathbb{Q}}^+$. Both $g = (1, m^{-1})$ and $\tilde m = (m, 1)$ are in $\mathcal{J}^+$ and have the same class in the idèle class group $C_{\mathbb{Q}}$, since $m\, g = \tilde m$. Thus the automorphisms $\vartheta_{\rm a}(g)$ and $\vartheta_{\rm a}(\tilde m)$ of $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$ are inner conjugate. Since the open set $\mathbb{A}_{\mathbb{K}}^{(1)} \subset \mathbb{A}_{\mathbb{K}}$ is not closed its characteristic function is not continuous and does not define a multiplier of $\mathcal{S}(\mathcal{G}_{\mathbb{K}})$. It follows that the endomorphism $F(g)$ is inner conjugate to the endomorphism $\vartheta_{\rm a}(\tilde m)$ only in the following weaker sense. There exists a sequence of elements u_n of $\mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$ such that for any $f \in \mathcal{S}(\mathcal{G}_{\mathbb{K}}^{(1)})$ with compact support

$$F(g)(f) = u_n\, \vartheta_{\rm a}(\tilde m)(f)\, u_n^*,$$

holds for all n large enough.

7.5.5. *Arithmetic subalgebra, Frobenius and monodromy.*

We now describe a natural algebra of coordinates $\mathcal{B}$ on $\Xi_{\mathbb{Q}}$. As we saw in Chapter 3, the BC system of Chapter 3 §4.1 ([**30**]), as well as its arithmetic generalizations of [**86**] and [**90**], have the important property that they come endowed with an arithmetic structure given by an arithmetic subalgebra. The general framework of endomotives developed in §2 shows a broad class of examples where a similar arithmetic structure is naturally present. We consider here the issue of extending the construction of the "rational subalgebra" of the BC system to the algebra $\mathcal{S}(\mathcal{G}_{\mathbb{Q}}^{(1)})$ of §7.5.4.

In order to get a good geometric picture it is convenient to think in terms of $\mathbb{Q}$-lattices rather than of adèles, as in Chapter 3 §4 ([**86**]). Thus, we use Lemma 3.21 of Chapter 3 to identify the set of 1-dimensional $\mathbb{Q}$-lattices with $\hat{\mathbb{Z}} \times \mathbb{R}_+^*$ by the map

$$\ell : \hat{\mathbb{Z}} \times \mathbb{R}_+^* \to \{\mathbb{Q} - \text{lattices}\}, \qquad \ell(\rho, \lambda) = (\Lambda, \phi) = (\lambda^{-1}\mathbb{Z}, \lambda^{-1}\rho), \tag{4.521}$$

which associates to an adèle $(\rho, \lambda) \in \hat{\mathbb{Z}} \times \mathbb{R}_+^* \subset \mathbb{A}_{\mathbb{Q}}$ the $\mathbb{Q}$-lattice obtained using ρ to label the torsion points of $\mathbb{R}/\lambda^{-1}\mathbb{Z}$. Replacing (ρ, λ) by $(n\rho, n\lambda)$, for a positive integer $n \in \mathbb{N}$, one obtains the pair $(\frac{1}{n}\Lambda, \phi)$, which is commensurable to (Λ, ϕ). Thus, the action of $\mathbb{Q}_+^*$ corresponds to commensurability of $\mathbb{Q}$-lattices under the map ι. Multiplying λ by a positive scalar corresponds to the scaling action of $\mathbb{R}_+^*$ on $\mathbb{Q}$-lattices.

Let us recall from Chapter 3, Theorem 3.30, the description of the "rational algebra" $\mathcal{A}_{1,\mathbb{Q}}$ for the BC system, in terms of $\mathbb{Q}$-lattices. We let, as in (3.83),

$$\epsilon_{1,a}(\Lambda, \phi) = \sum_{y \in \Lambda + \phi(a)} y^{-1}, \tag{4.522}$$

for any $a \in \mathbb{Q}/\mathbb{Z}$. This is well defined, for $\phi(a) \neq 0$, using the summation $\lim_{N\to\infty} \sum_{-N}^{N}$, and is zero by definition for $\phi(a) = 0$. The function

$$\varphi_a(\rho, \lambda) = \epsilon_{1,a}(\ell(\rho, \lambda)), \quad \forall (\rho, \lambda) \in \hat{\mathbb{Z}} \times \mathbb{R}_+^*, \tag{4.523}$$

is well defined and homogeneous of degree 1 in λ. Moreover, for fixed $a \in \mathbb{Q}/\mathbb{Z}$ with denominator m, it only depends upon the projection of ρ on the finite group $\mathbb{Z}/m\mathbb{Z}$, hence it defines a continuous function on $\hat{\mathbb{Z}} \times \mathbb{R}_+^*$. Using the degree 1 homogeneity in λ, one gets that (4.523) extends by continuity to 0 on $\hat{\mathbb{Z}} \times \{0\}$.

One gets functions that are homogeneous of degree zero by taking the derivatives of the functions φ_a. The functions

$$\psi_a(\rho, \lambda) = \frac{1}{2\pi i}\frac{d}{d\lambda}\varphi_a(\rho, \lambda), \quad \forall (\rho, \lambda) \in \hat{\mathbb{Z}} \times \mathbb{R}_+^*, \tag{4.524}$$

are independent of λ, hence they define continuous functions on $\mathbb{A}_{\mathbb{Q}}^{(1)}$. They are non-trivial on $\tilde{\Xi}_{\mathbb{Q},\infty} = \hat{\mathbb{Z}}^* \times \{0\} \subset \hat{\mathbb{Z}} \times \{0\}$ and they agree there with the functions $e_{1,a}$ of Chapter 3 (3.80).

Proposition 4.181. *Let $\mathcal{B}$ be the algebra generated by the φ_a and ψ_a defined in* (4.523) *and* (4.524) *above.*

(1) *The expression*

$$N(f) = \frac{1}{2\pi i}\frac{d}{d\lambda} f \tag{4.525}$$

defines a derivation N of $\mathcal{B}$.

(2) *The algebra $\mathcal{B}$ is stable under the derivation Y that generates the* 1*-parameter semigroup $F(\mu)$ of endomorphisms of $\mathcal{S}(\mathcal{G}_{\mathbb{Q}}^{(1)})$ of* (4.520) *and one has, at the formal level, the relation*

$$F(\mu)N = \mu N F(\mu). \tag{4.526}$$

(3) *For any element $f \in \mathcal{B}$ one has*

$$\alpha \circ f(x) = f(\tilde{\alpha}\, x), \quad \forall x \in \tilde{\Xi}_{\mathbb{Q},\infty} \text{ and } \forall \alpha \in \mathrm{Gal}(\mathbb{Q}^{cycl}/\mathbb{Q}), \tag{4.527}$$

where $\tilde{\alpha} \in \hat{\mathbb{Z}}^* \subset C_{\mathbb{Q}}$ is the element of the idèle class group associated to $\alpha \in \mathrm{Gal}(\mathbb{Q}^{cycl}/\mathbb{Q})$ by the class field theory isomorphism.

PROOF. (1) By construction N is a derivation of the algebra of functions. Moreover (4.524) shows that $N(\varphi_a) = \psi_a$, while $N(\psi_a) = 0$. Thus, the derivation rule shows that $\mathcal{B}$ is stable under N.
(2) The derivation generating the one-parameter semigroup $F(\mu)$ is given, up to sign, by the grading operator

$$Y(f) = \lambda \frac{d}{d\lambda} f. \tag{4.528}$$

By construction, any of the φ_a is of degree one, i.e. $Y(\varphi_a) = \varphi_a$ and ψ_a is of degree 0. Thus, again the derivation rule shows that $\mathcal{B}$ is stable under Y.
(3) This only involves the functions ψ_a, since by construction the restriction of φ_a is zero on $\tilde{\Xi}_{\mathbb{Q},\infty}$. The result then follows from Chapter 3, (3.137) in Theorem 3.32. In fact, all these functions take values in the cyclotomic field $\mathbb{Q}^{cycl} \subset \mathbb{C}$ and they intertwine the action of the discontinuous piece $\hat{\mathbb{Z}}^*$ of $C_{\mathbb{Q}}$ with the action of the Galois group of $\mathbb{Q}^{cycl}$. □

This is in agreement with viewing the algebra $\mathcal{B}$ as the algebra of coordinates on $\tilde{\Xi}_{\mathbb{Q}}$. Indeed, in the case of a global field $\mathbb{K}$ of positive characteristic, the action of the Frobenius on the points of $C(\bar{\mathbb{F}}_q)$ (which have coordinates in $\bar{\mathbb{F}}_q$) corresponds to the Frobenius map

$$\mathrm{Fr} : u \mapsto u^q, \qquad \forall u \in \mathbb{K} \tag{4.529}$$

of the function field $\mathbb{K}$ of the curve C. The Frobenius endomorphism $u \mapsto u^q$ of $\mathbb{K}$ is the operation that replaces a function $f : C(\bar{\mathbb{F}}_q) \to \bar{\mathbb{F}}_q$ by its q-th power, i.e. the composition $\mathrm{Fr} \circ f$ with the Frobenius automorphism $\mathrm{Fr} \in \mathrm{Gal}(\bar{\mathbb{F}}_q/\mathbb{F}_q)$. For $f \in \mathbb{K}$, one has

$$\mathrm{Fr} \circ f = f^q = f \circ \mathrm{Fr}, \tag{4.530}$$

where on the right hand side Fr is the map that raises every coordinate to the power q. This corresponds to the intertwining with the Galois action discussed above.
Notice moreover that, as we have seen in Proposition 4.179, only the semigroup $C_{\mathbb{Q}}^+$ acts on the reduced system $\mathcal{S}(\mathcal{G}_{\mathbb{Q}}^{(1)})$ and it acts by endomorphisms. It nevertheless acts in a bijective manner on the points of $\Xi_{\mathbb{Q}}$. This is similar to what happens with the Frobenius endomorphism (4.529), which is only an endomorphism of the field of functions $\bar{\mathbb{K}}$, while it acts bijectively (as a Galois automorphism of the coordinates) on the points of $C(\bar{\mathbb{F}}_q)$.

Further notice that there is a striking formal analogy between the operators F and N of Proposition 4.181 satisfying the relation (4.526) and the Frobenius and local monodromy operators introduced in the context of the "special fiber at arithmetic infinity" in Arakelov geometry (see [**100**], [**102**]). In particular, one should compare (4.526) with §2.5 of [**102**] that discusses a notion of Weil–Deligne group at arithmetic infinity.

7.6. Vortex configurations (an analogy).

We formulate here an analogy between some of the beginning steps of the Weil proof of RH for function fields and statements formulated in terms of differential operators and index theory. The purpose is to help in building a dictionary between correspondences given in terms of divisors in the Weil proof and correspondences in terms of something akin to Kasparov modules and (pseudo)differential operators in noncommutative geometry.

Since the purpose of this section is mainly to build an analogy, we will not give the full analytic details of the vortex moduli spaces we consider below, but we include references to the literature where the reader can find a more detailed treatment.

We assume that C is a smooth algebraic curve and we think of it as a curve over $\mathbb{C}$ (a compact Riemann surface) to obtain the differential geometric analogs of the arguments given by Weil in the function field case (for a curve defined over a finite field). It is already a well known fact that the complex geometry analog of positivity can be found in the Hodge index theorem, in the form of Castelnuovo positivity (cf. e.g. [**266**]).

As in §6.2 above, we consider a correspondence, given by a divisor Z on $C \times C$. As we have seen, this can be viewed as a multivalued map (4.388). We consider the two degrees (4.383).

As we have seen in §6 above, in the Weil argument one first uses trivial correspondences $P \times C$ and $C \times P$ to alter the degree of $Z(P)$. Thus, up to trivial correspondences, we can start by assuming that we are handed a correspondence $Z(P)$ with

$$d(Z) = g, \tag{4.531}$$

where g is the genus of the curve.

We begin by translating these data in more operator theoretic language. For a fixed $P \in C$, we are dealing with a divisor $D = Z(P)$ on the Riemann surface C. There is an associated line bundle $L = L_D$ on C.

LEMMA 4.182. *The Riemann–Roch formula for the curve C can be equivalently written as the Atiyah–Singer index theorem*

$$\operatorname{Ind} \bar{\partial}_L = \int_C \operatorname{ch}(L) \wedge \operatorname{td}(TC). \tag{4.532}$$

PROOF. The right hand side of (4.532) can be computed as

$$\begin{aligned} \textstyle\int_C \operatorname{ch}(L) \wedge \operatorname{td}(TC) = &\ \textstyle\int_C (1 + c_1(L))(1 + \frac{1}{2} c_1(TC)) \\ = \textstyle\int_C c_1(L) + \frac{1}{2}\int_C c_1(TC) = &\ \deg(L) + 1 - g. \end{aligned} \tag{4.533}$$

This recovers the usual Riemann-Roch formula upon identifying

$$\operatorname{Ker} \bar{\partial}_L = H^0(C, L) \tag{4.534}$$

and

$$\operatorname{Coker} \bar{\partial}_L = \operatorname{Ker} \bar{\partial}_L^* = H^1(C, L) = H^0(C, \mathcal{K} \otimes L^{-1}), \tag{4.535}$$

where $\mathcal{K} \simeq T^*C$ is the canonical line bundle and the last identification is given by Serre duality. ☐

One sees from the formulae (4.532) an (4.533) that if we consider a line bundle L with $\deg(L) = g$, which is what we obtain from (4.531), then we have $\operatorname{Ind}\partial_L = 1$, which guarantees that $\dim \operatorname{Ker}\partial_L > 0$.

There is another natural differential operator on the Riemann surface C, which differs slightly from the $\bar\partial$ operator considered above. It is the Dirac operator $\not\partial$. On a compact Riemann surface C this is related to the $\bar\partial$ operator by

$$\not\partial = \bar\partial_{\mathcal{K}^{1/2}}. \tag{4.536}$$

As in the case of the $\bar\partial$ operator, one can consider the Dirac operator twisted with a line bundle L,

$$\not\partial_L = \bar\partial_{\mathcal{K}^{1/2}\otimes L}. \tag{4.537}$$

The Atiyah-Singer index theorem for $\not\partial_L$ is then of the form

$$\operatorname{Ind}\not\partial_L = \int_C \operatorname{ch}(L)\wedge \hat A(TC) = \deg(L). \tag{4.538}$$

Thus, in the case of the Dirac operator, if we consider a line bundle L with $\deg(L) = g$ we obtain not just a positive dimensional kernel, but the stronger condition $\operatorname{Ind}\not\partial_L = g$, which implies $\dim\operatorname{Ker}\not\partial_L > g-1$. We will encounter below another differential operator that has this same property.

In the Weil proof one then proceeds to modify the divisor $Z(P)$ to an effective divisor, while maintaining the same degree $d(Z) = g$. This is done using the fact that any divisor of degree g is linearly equivalent to an effective divisor. We want to see the same result, and the effect of moving $D = Z(P)$ by a principal divisor, in operator theoretic terms.

We first recall the following procedure, which is standard when dealing with the index for families (cf. e.g. §5.1.3 and §5.2.1 of [**121**]). Notice that we are in fact dealing with a family case, since the divisor $D = Z(P)$ depends upon the point $P \in C$, through the correspondence Z, so that we are in fact dealing with a family of differential operators

$$\mathcal{D}_P = \not\partial_{L_{Z(P)}} \quad \text{or} \quad \mathcal{D}_P = \bar\partial_{L_{Z(P)}}, \tag{4.539}$$

parameterized by $P \in C$.

Suppose given an elliptic differential operator $\mathcal{D}$ with pointwise index satisfying $\operatorname{Ind}\mathcal{D} > 0$. For the moment, we can assume that $\mathcal{D}$ is either $\bar\partial_L$ or $\not\partial_L$, with the line bundle L associated to the divisor $D = Z(P)$ as in (4.539), but we will soon focus on a slightly different choice of $\mathcal{D}$. The operator acts between smooth sections of vector bundles,

$$\mathcal{D} : \Gamma(E) \to \Gamma(E').$$

The formal difference $[\operatorname{Ker}\mathcal{D}_P] - [\operatorname{Coker}\mathcal{D}_P]$ defines an element in the K-theory of C. We use the notation $\mathcal{E}_P = \Gamma(C, E_P)$ and $\mathcal{E}'_P = \Gamma(C, E'_P)$. The map $\mathcal{D}_P$ is Fredholm, upon fixing suitable Sobolev completions of the spaces $\mathcal{E}_P$ and $\mathcal{E}'_P$. If $\mathcal{D}$ is surjective, then $\operatorname{Ker}\mathcal{D}$ is a vector bundle over C. Otherwise, one can find a trivial vector bundle $\mathbb{C}^N$ on C and a map $T : \mathbb{C}^N \to \mathcal{E}'_P$ such that the linear operator $\mathcal{D}\oplus T$ is surjective, and one has

$$\operatorname{Ind}(\mathcal{D}) = [\operatorname{Ker}(\mathcal{D}\oplus T)] - [\mathbb{C}^N] \in K(C).$$

The map T is defined locally so that its image spans the $\mathrm{Coker}\,\mathcal{D}_P$ and patched together globally. The index class does not depend on the choice of such map T. The determinant line bundle of $\mathcal{D}$ is then defined equivalently as

$$\begin{aligned} \det \mathcal{D} &= \bigwedge^{\max} \mathrm{Ker}\,\mathcal{D} \otimes (\bigwedge^{\max} \mathrm{Coker}\,\mathcal{D})^* \\ &= \bigwedge^{\max} \mathrm{Ker}\,(\mathcal{D} \oplus T) \otimes (\bigwedge^{N} \mathbb{C}^N)^*. \end{aligned} \tag{4.540}$$

The right hand side shows that indeed $\det \mathcal{D}$ is a well-defined line bundle, regardless of the fact that the dimensions of $\mathrm{Ker}\,\mathcal{D}$ and $\mathrm{Coker}\,\mathcal{D}$ may jump.

As observed in §5.2.1 of [**121**], passing from $\mathcal{D}$ to $\mathcal{D} \oplus T$ by the construction above is closely related to the procedure of transverse perturbations, widely used in gauge theory to achieve transversality of the relevant moduli spaces (cf. e.g. §4.2 of [**121**]).

In that setting, typically, the virtual dimension of a moduli space is computed by an index theorem for the Fredholm operator given by the linearization at a solution of a system of nonlinear elliptic differential equations. The value of the index gives the virtual (expected) dimension of the moduli space. In fact, while the kernel of the operator represents infinitesimal deformations of the solutions (in the Kuranishi model), the cokernel represents obstructions to the smoothness and transversality of the moduli space.

We now rephrase in these "gauge theoretic" terms the operation of modifying the divisor $D = Z(P)$ by a principal divisor. To this purpose we rephrase the data of divisors in terms of *vortex configurations*.

Definition 4.183. *For a given holomorphic line bundle L over a compact Riemann surface C, one considers the non-linear Yang–Mills functional*

$$E(A,u) = \frac{1}{2}\|F_A\|^2 + 2\left\|\frac{1}{1+|u|^2}\,\bar\partial_A u\right\|^2 + \frac{1}{2}\left\|\frac{1-|u|^2}{1+|u|^2}\right\|^2, \tag{4.541}$$

where A is a $U(1)$-connection on L and u is a section.

One then has the following result.

Lemma 4.184. *The first order equations associated to the variational problem that minimizes the energy functional give rise to the self-dual (or anti-self-dual) equations of abelian Higgs theory,*

$$\bar\partial_A u = 0, \qquad *iF_A = \frac{1-|u|^2}{1+|u|^2}. \tag{4.542}$$

The moduli space $\mathcal{M}_L(C)$ of solutions to the equations (4.542) *modulo gauge transformations $Maps(C, U(1))$ is a copy of the symmetric product of the curve C,*

$$\mathcal{M}_L(C) = S^d(C), \tag{4.543}$$

where $d = \deg(L)$.

The moduli space (4.543) corresponds to the effective divisors of holomorphic sections of the line bundle L,

$$D = u^{-1}(0) = (Q_1, \ldots, Q_d) \in S^d(C).$$

The identification (4.543) actually holds only in a certain range of values of $\deg(L)$. To extend the range of validity to arbitrary $\deg(L)$ one can introduce an additional sufficiently large real parameter in the equation, which corresponds to scaling the volume, as in [**32**]. We discuss this point later.

The configuration space $\mathcal{B}_L(C)$ is set of all pairs (A, u) of connections and smooth sections of L modulo gauge, completed in suitable Sobolev norms so that the linearization of the equations is a Fredholm operator. The moduli space (4.543) is cut out transversely by the equations (4.542) inside the configuration space $\mathcal{B}_L(C)$. This means that the linearization, which gives the deformation complex for the moduli space, has trivial cokernel. This follows, for instance, from the results of [**137**].

There is a way to extend the gauge theoretic picture to include the case of meromorphic sections (and non-effective divisors). The functional (4.541) extends to a Yang–Mills functional defined on sections of the complex projective bundle $\mathbb{P}(L \oplus \mathbb{C})$ over C (cf. [**270**], [**209**]) and this corresponds to a gauge theory on Riemann surfaces where vortices and anti-vortices simultaneously arise, namely, where the section u is a meromorphic section with a certain number n of poles. Since, if m denotes the number of zeros, the difference $m - n = \int_C c_1(L) = \deg(L)$ is fixed, assigning n also fixes m.

We need to define a suitable configuration space $\mathcal{B}_L(C, n)$ for these data, with $\mathcal{B}_L(C) = \mathcal{B}_L(C, 0)$. It is convenient to consider the elements of $\mathcal{B}_L(C, n)$ as gauge equivalence classes of pairs $(\tilde{A}, \tilde{u})$, where $\tilde{A}$ is a $U(1)$-connection and $\tilde{u}$ is a smooth section of the projectivized bundle $\mathbb{P}(L \oplus \mathbb{C})$. The numbers of zeros and poles are then specified by the data

$$m = \#(\tilde{u}(C) \cap s_0(C)), \quad n = \#(\tilde{u}(C) \cap s_\infty(C)),$$

where s_0 and s_∞ are the two natural sections of $\mathbb{P}(L \oplus \mathbb{C})$,

$$s_0(x) = \ell_{(0_x,1)}, \quad s_\infty(x) = \ell_{(v_x,0)},$$

where $\ell_{(v_x,u)}$ denotes the complex line through the point $(v_x, u) \in L_x \oplus \mathbb{C}$.

In this way, one introduces two different charts,

$$\Omega_\infty = \{\tilde{u} \cap s_0 = \emptyset\}, \quad \Omega_0 = \{\tilde{u} \cap s_\infty = \emptyset\}.$$

In the chart Ω_0 the section $\tilde{u}$ can be identified with a section of L, while in the chart Ω_∞, $\tilde{u}$ can be identified with a section $u^\vee$ of the dual bundle $\mathrm{Hom}(L, \mathbb{C})$. Here we view $\mathrm{Hom}(L, \mathbb{C}) \subset \mathbb{P}(L \oplus \mathbb{C})$ by identifying $h \in \mathrm{Hom}(L, \mathbb{C})$ with the complex line $\{(v_x, h(v_x)) : x \in C\}$. The connection can also be identified with a connection $A^\vee$ on the dual line bundle.

LEMMA 4.185. *The functional* (4.541) *then extends to all of* $\mathcal{B}_L(C, n)$, *for any given* n, *in the form*

$$E(\tilde{A}, \tilde{u}) = \begin{cases} \frac{1}{2}\|F_A\|^2 + 2\left\|\frac{1}{1+|u|^2}\,\bar{\partial}_A u\right\|^2 + \frac{1}{2}\left\|\frac{1-|u|^2}{1+|u|^2}\right\|^2, & \tilde{u} \in \Omega_0 \\ \frac{1}{2}\|F_{A^\vee}\|^2 + 2\left\|\frac{1}{1+|u^\vee|^2}\,\bar{\partial}_{A^\vee} u^\vee\right\|^2 + \frac{1}{2}\left\|\frac{1-|u^\vee|^2}{1+|u^\vee|^2}\right\|^2, & \tilde{u} \in \Omega_\infty. \end{cases} \tag{4.544}$$

The values of the functional agree in the overlap of the two charts.

The corresponding first order equations can also be seen in the two charts and they give the Bogomolny equations (4.542). These determine a meromorphic section u of L and a meromorphic section $u^\vee$ of the dual line bundle. The zeros $\tilde{u}(C) \cap s_0(C)$ are the zeros of u and the poles $\tilde{u}(C) \cap s_\infty(C)$ are the zeros of $u^\vee$.

A global solution $(\tilde{A}, \tilde{u})$ thus determines local solutions (A, u) and $(A^\vee, u^\vee)$ of (4.542) in Ω_0 and Ω_∞, with u and $u^\vee$ meromorphic.

The reason for going through this gauge theoretic reformulation of the data of a holomorphic line bundle L with a Hermitian connection A and a meromorphic section is that now we can give a reformulation of the operation of modifying a divisor D by a principal divisor so as to obtain an effective divisor in terms of deforming a first order elliptic differential operator by a zero order perturbation to make it surjective.

The family of differential operators we consider here are given by the linearization of (4.542) at a solution.

LEMMA 4.186. *The linearization of the equations* (4.542) *at a solution* (A, u) *give a first order elliptic differential operator* $\mathcal{D}_{(A,u)}$ *acting on pairs* (α, w) *of a 1-form* $\alpha = \alpha^{0,1} + \alpha^{1,0}$ *and a section* w*. This is explicitly of the form*

$$\mathcal{D}_{(A,u)}(\alpha, w) = \begin{cases} \bar\partial_A w + i\alpha^{0,1} \cdot u \\ - * i d\alpha + \ell(|u|^2)\Re\langle u, w\rangle \\ d^*\alpha - i\Im\langle u, w\rangle, \end{cases} \tag{4.545}$$

where, in the second equation we have

$$\ell(x) = 2\left(x^2 - \frac{1}{1+x^2}\right).$$

The third equation is the gauge fixing condition (the slice of the gauge action) given by the L^2-adjoint $G^*_{(A,u)}(\alpha, w)$ of the infinitesimal gauge action

$$G_{(A,u)}(f) = (-idf, if\, u).$$

We obtain the index of the linearization (4.545).

LEMMA 4.187. *The index of* $\mathcal{D}_{(A,u)}$ *is given by*

$$\mathrm{Ind}\,\mathcal{D}_{(A,u)} = \mathrm{Ind}\,\bar\partial_L + \mathrm{Ind}\,\bar\partial^* = \deg(L). \tag{4.546}$$

PROOF. This follows from the identity $d^*\alpha - i * d\alpha = 2\bar\partial\alpha^{0,1}$. In fact, this shows that, up to zero order terms, one can compute $\mathrm{Ind}\,\mathcal{D}_{(A,u)}$ as $\mathrm{Ind}\,\bar\partial_L + \mathrm{Ind}\,\bar\partial^*$. $\square$

COROLLARY 4.188. *If* (A, u) *and* $(A^\vee, u^\vee)$ *are the data coming from the restriction of* $(\tilde{A}, \tilde{u})$ *to the charts* Ω_0 *and* Ω_∞*, respectively, then*

$$\mathrm{Ind}\,\mathcal{D}_{(A^\vee,u^\vee)} = \mathrm{Ind}\,\bar\partial_{L^{-1}} + \mathrm{Ind}\,\bar\partial^* = -\mathrm{Ind}\,\mathcal{D}_{(A,u)}. \tag{4.547}$$

Lemma 4.187 gives the virtual (complex) dimension of the space of gauge classes of solutions (A, u) of the equations (4.542). If the linearization (4.545) is surjective, then the virtual dimension equals the actual dimension, while if the dimension of the kernel is larger than the index, there may be excess intersection and the moduli space will not be cut out transversely by the equations.

As we remarked before, the virtual dimension being positive does not guarantee that the moduli space will be nonempty. Conversely, there may be cases where the moduli space is nonempty though having a negative virtual dimension. Thus, computing the virtual dimension does not suffice and we need to find the actual dimension of the moduli spaces $\mathcal{M}_L(C, n)$.

It is in fact known what the actual dimension of the moduli spaces $\mathcal{M}_L(C,n)$ is. First one needs to ensure that the $\mathcal{M}_L(C,n)$ are nonempty. A necessary condition is seen if one integrates the curvature equation of (4.542).

LEMMA 4.189. *A necessary condition for the equation* (4.542) *to admit solutions with nontrivial section* u *is the estimate*

$$|\deg(L)| < \frac{\mathrm{Vol}(C)}{2\pi}. \tag{4.548}$$

PROOF. The term with the curvature form yields the topological invariant $2\pi\deg(L)$ and the term $1/(1+|u|^2)$ is bounded above by 1. Thus one obtains an estimate of the form (4.548). □

This bound can be improved upon inserting a positive real parameter τ in the equation (replacing $1-|u|^2$ with $\tau-|u|^2$). This relaxes the condition (4.548) to the weaker

$$|\deg(L)| < \frac{\tau\mathrm{Vol}(C)}{2\pi}. \tag{4.549}$$

The main result of [**270**] shows the following.

PROPOSITION 4.190. *For an assigned divisor*

$$D = \{q_1,\dots,q_m,-p_1,\dots,-p_n\} \tag{4.550}$$

of degree

$$\deg(D) = m-n = \deg(L),$$

there exists a unique solution $(\tilde{A},\tilde{u})$ *(in the gauge slice).*

The argument is given as a fixed point argument, as is customary in problems of pasting of solutions of nonlinear elliptic equations. We do not reproduce a proof here. The result is proved in [**270**] for the unperturbed equations ($\tau=1$) with the constraint (4.548), while the fixed point argument for a perturbed version of the equations is considered in [**209**], although with a different choice of perturbation and a variant of (4.549).

As a consequence of Proposition 4.190 one can also obtain the actual dimension of the moduli space of solutions.

COROLLARY 4.191. *The actual dimension of the moduli space* $\mathcal{M}_L(C,n)$ *is*

$$\dim_{\mathbb{C}} \mathcal{M}_L(C,n) = n+m. \tag{4.551}$$

PROOF. Upon describing solutions in terms of the data of divisors as in Proposition 4.190, one sees that the freedom to move along the moduli space from a point $[\tilde{A},\tilde{u}]$ corresponds to the freedom to move the choice of the $n+m$ points (zeros and poles) of the divisor D. □

We can now formulate the following result.

PROPOSITION 4.192. *Suppose given a divisor* D *on the Riemann surface* C *of assigned positive degree and a continuous family of divisors* D_t*, depending on a real parameter* $t\in[0,1]$*, such that* $D_0 = D = \{q_1,\dots,q_m,-p_1,\dots,-p_n\}$ *and with* D_1 *effective. There is a corresponding family* $\mathcal{D}_{(A_t,u_t)}$ *of first order elliptic differential operators with* $\mathrm{Ind}\,\mathcal{D}_{(A_t,u_t)} = \deg(D_t) = \deg(D)$ *and such that* $\dim\mathrm{Coker}\,\mathcal{D}_{(A_0,u_0)} = n$ *and* $\dim\mathrm{Coker}\,\mathcal{D}_{(A_1,u_1)} = 0$.

PROOF. If one considers the restriction (A,u) of a solution $(\tilde{A},\tilde{u})$ to Ω_0, thus freezing the n modes corresponding to poles, one obtains then that $\dim \operatorname{Ker} \mathcal{D}_{(A,u)} = m$ (the number of zeros) and, from the index computation (4.546), $\dim \operatorname{Coker} \mathcal{D}_{(A,u)} = n$ (the number of poles).

A neighborhood of the solution $(\tilde{A},\tilde{u})$ in the moduli space $\mathcal{M}_L(C,n)$ can be obtained via a gluing construction. Upon writing the divisor D of (4.550) in the form $D = D_0 - D_\infty$ with

$$D_0 = \{q_1,\dots,q_m\} \quad D_\infty = \{p_1,\dots,p_n\}, \tag{4.552}$$

one can consider the moduli space $\mathcal{M}_L(C,D_\infty)$ of gauge classes of solutions (A,u) of the equations (4.542), where u is a meromorphic section with poles at D_∞. Similarly, one considers the moduli space $\mathcal{M}_{L^{-1}}(C,D_0)$ of gauge classes of solutions $(A^\vee,u^\vee)$ of the equations (4.542) on the dual line bundle L^{-1}, with $u^\vee$ a meromorphic section with poles at D_0. One can paste together elements of $\mathcal{M}_L(C,D_\infty)$ and $\mathcal{M}_{L^{-1}}(C,D_0)$ by a cutoff function and create approximate solutions

$$(A,u)\#(A^\vee,u^\vee) \tag{4.553}$$

in a neighborhood $\mathcal{U}_{\tilde{A},\tilde{u})}$ of the original solution $(\tilde{A},\tilde{u})$ in the configuration space $\mathcal{B}_L(C,n)$. The dimension of this neighborhood is given by the "approximate kernel" of the linearization (4.545) at the approximate solution, that is, by the span of the eigenvectors with small eigenvalues of the operator

$$\mathcal{D}^*_{(A,u)\#(A^\vee,u^\vee)}\mathcal{D}_{(A,u)\#(A^\vee,u^\vee)}. \tag{4.554}$$

This span can be seen to have dimension

$$\dim \operatorname{Ker} \mathcal{D}_{(A,u)} + \dim \operatorname{Ker} \mathcal{D}_{(A^\vee,u^\vee)}. \tag{4.555}$$

Inside this neighborhood of the solution $(\tilde{A},\tilde{u})$ one finds the neighborhood of actual solutions

$$\mathcal{V}_{(\tilde{A},\tilde{u})} = \mathcal{U}_{(\tilde{A},\tilde{u})} \cap \mathcal{M}_L(C,n) \tag{4.556}$$

as the zero set of an obstruction section

$$\mathcal{V}_{(\tilde{A},\tilde{u})} = \sigma^{-1}(0) \subset \mathcal{U}_{(\tilde{A},\tilde{u})} \tag{4.557}$$

of a bundle over $\mathcal{U}_{(\tilde{A},\tilde{u})}$ with fiber the approximate cokernel, i.e. the small eigenmodes of

$$\mathcal{D}_{(A,u)\#(A^\vee,u^\vee)}\mathcal{D}^*_{(A,u)\#(A^\vee,u^\vee)}),$$

of dimension

$$\dim \operatorname{Coker} \mathcal{D}_{(A,u)} + \dim \operatorname{Coker} \mathcal{D}_{(A^\vee,u^\vee)}. \tag{4.558}$$

The fixed point argument of [**270**] translates into the fact that the obstruction section vanishes and all approximate solutions can be perturbed to actual solutions (by an iteration method). Thus, the $n+m$-dimensional tangent space of the moduli space $\mathcal{M}_L(C,n)$ can be identified with

$$\mathcal{T}_{(\tilde{A},\tilde{u})}\mathcal{M}_L(C,n) \simeq \operatorname{Ker}\mathcal{D}_{(A,u)} \times \operatorname{Ker}\mathcal{D}_{(A^\vee,u^\vee)}. \tag{4.559}$$

Suppose now given a continuous family of divisors D_t, depending on a real parameter $t \in [0,1]$, such that $D_0 = D = \{q_1,\dots,q_m,-p_1,\dots,-p_n\}$ and with D_1 effective. We have

$$\deg(D_t) = \deg(D_0) = m - n = \deg(L).$$

By the result on existence and uniqueness of solutions described above, this family D_t determines a path $[\tilde{A}_t, \tilde{u}_t]$. There will be intervals $[t_k, t_{k+1}]$ with $t_0 = 0$ and $t_N = 1$ such that, for all $t \in [t_k, t_{k+1}]$, we have $[\tilde{A}_t, \tilde{u}_t] \in \mathcal{M}_L(C, n_k)$, for some $n_k \geq 0$, with $[\tilde{A}_0, \tilde{u}_0] \in \mathcal{M}_L(C, n)$ and $[\tilde{A}_1, \tilde{u}_1] \in \mathcal{M}_L(C, 0) = s^{\deg(L)}(C)$.

If we consider the corresponding one-parameter family of linear differential operators $\mathcal{D}_{(A_t,u_t)}$ we see that

$$\dim \operatorname{Coker} \mathcal{D}_{(A_0,u_0)} = n \quad \text{while} \quad \dim \operatorname{Coker} \mathcal{D}_{(A_1,u_1)} = 0.$$

Thus, we have shown that deforming a given divisor D of assigned (positive) degree to an effective divisor of the same degree corresponds to deforming the differential operator $\mathcal{D}_{(A,u)}$ to a surjective operator while preserving the index. □

7.7. Building a dictionary.

We have described some of the first steps towards building a dictionary between the Weil proof of RH for function fields and the noncommutative geometry of the adèle class space. We outline here some of the entries of the dictionary, some of which we have already described and some that are still more tentative. They are summarized in Table 1 below.

Correspondences in the noncommutative geometry setting, as we have seen at the beginning of this chapter, can be thought of in terms of Kasparov modules and KK-theory classes, with the Kasparov product behaving like composition of correspondences. As we already pointed out in §1, the equivalence relation of homotopy imposed on Kasparov modules to obtain KK-theory classes is most likely not the best equivalence relation to work with when one wants a good analog of correspondences given by algebraic cycles. In the zero-dimensional setting of endomotives this is not a problem as the dependence on the equivalence relation is no longer seen, as in the case of Artin motives, cf. §2, but more generally this issue will have to be addressed.

We have seen that we do have a natural replacement for the Frobenius correspondences $\sum_n a_n \mathrm{Fr}^n$, given by the scaling action of $C_{\mathbb{K}}$, in the form of the elements

$$Z(f) = \int_{C_{\mathbb{K}}} f(g)\, Z_g \, d^*g \quad \text{with} \quad f \in \mathbf{S}\,(C_{\mathbb{K}}).$$

There is also a good analog for the degree of a correspondence, which should be identified with a pointwise index, if we think of the correspondence $Z(f)$ as giving rise to a bivariant class $\Gamma(f)$. (This is a necessary step of quantization of the Frobenius correspondence.) We have also seen that it is possible to modify the degree of a correspondence $Z(f)$ by adding to f an element in the range $\mathcal{V}$ of the restriction map $\mathrm{Tr} \circ \pi$.

The next very important step is the perturbation of a correspondence Z with degree $\deg(Z) = g$ to an effective correspondence, obtained by adding a principal divisor $\mathrm{div}(f)$. This lacks an immediate analog. We have given in §7.6 above a heuristic argument which suggests that the analog of changing a correspondence by principal divisors to make it effective can be thought of as achieving surjectivity of a Fredholm operator of positive index by a compact perturbation that yields an operator with

(Virtual) correspondences	Bivariant class Γ
Composition	Cup product in KK-theory
Frobenius correspondence	$Z(f) \Rightarrow$ bivariant element $\Gamma = \Gamma(f)$
Degree of correspondence	Pointwise index $d(\Gamma)$
Trivial correspondences	Elements of the range $\mathcal{V}$
Adjusting the degree by trivial correspondences	Failure of Fubini on the test functions
$\ell(D) - \ell(K - D) = \deg D - (g - 1)$	$\mathrm{Ind}\, \bar{\partial}_L = \mathrm{td}(TC) + c_1(L)$
$\deg Z(P) \geq g \Rightarrow Z + \mathrm{div}(f)$ effective	$d(\Gamma) > 0 \Rightarrow \Gamma + K$ onto
Effective correspondences	Epimorphism of C^*-modules
Modulo torsion	$KK(A, B \otimes \mathrm{II}_1)$
Lefschetz formula	Bivariant Chern of $\Gamma(f)$ (by localization on the graph $Z(f)$)
Weil trace unchanged by principal divisors	Bivariant Chern unchanged by compact perturbations

TABLE 1. The Weil proof and the adèle class space: a dictionary

the same index and a trival cokernel. The Riemann–Roch formula becomes in this context the index theorem. This in particular requires interpreting the degree of a correspondence $Z(f)$ as the local index of an associated bivariant class $\Gamma(f)$.

In Weil's proof the positivity of the trace ignores all the torsion in the algebra of correspondences under composition. Thus, in building an appropriate dictionary, one should perform a similar simplification. In the context of KK-theory, replacing $KK(\mathcal{A},\mathcal{B})$ with $KK(\mathcal{A},\mathcal{B}\otimes\mathcal{N})$, where $\mathcal{N}$ is a type II_1 factor, has the effect of removing torsion by tensoring the group with $\mathbb{R}$. This makes it possible to consider combinations of correpondences with *real* coefficients, due to the fact that dimension theory for II_1 factors gives $K(\mathcal{N})=\mathbb{R}$. Thus, to be able to work modulo torsion, it is best to work with correspondences obtained after tensoring with a type II_1 factor.

This discussion only touched upon some of the issues involved in the making of such a dictionary. The development of this project will also imply further developing and refining the tools of noncommutative geometry.

8. The analogy between QG and RH

We end this chapter and the book by returning to the problem of constructing a theory of quantum gravity. We use the understanding we gained in the last chapters about phase transitions in the quantum statistical mechanical systems of $\mathbb{Q}$-lattices and the analogy with the way the electroweak phase transition is treated in [**267**]. This suggests that the KMS condition may be taken as a definition of temperature states on a suitable algebra $\mathcal{H}_G$ of quantum gravity observables.

8.1. KMS states and the electroweak phase transition.

We discuss the electroweak phase transition in the Standard Model following [**267**]. It is based on the computation of the effective potential for the scalar Higgs field in the Standard Model and the effect of high temperature on symmetry breaking.

8.1.1. *Effective potential at one loop.* The effective action of Chapter I §3.5 can be expanded in powers of the derivatives of the classical field ϕ around the constant field $\phi=\phi_c$. This expansion, written in increasing powers of the momenta, is then of the form

$$S_{eff}(\phi)=\int[-V(\phi)+\frac{1}{2}(\partial_\mu\phi)^2Z(\phi)+\cdots]\,d^Dx \tag{4.560}$$

where V and Z are ordinary functions of one variable.

The function V is the *effective potential* and determines spontaneous symmetry breaking through minimization of V and the existence of a non-zero solution of

$$dV/d\phi_c=0\,,\quad \phi_c\neq 0 \tag{4.561}$$

Since the effective action is, in the Euclidean context, the formal power series in $\hbar$

$$S_{eff}(\phi)=\,S(\phi)-\sum_{\Gamma\in\,1\mathrm{PI}}\hbar^L\frac{\Gamma(\phi)}{\sigma(\Gamma)},$$

one finds that the effective potential V can be expressed as the following sum over 1PI diagrams with zero external momenta:

$$V(\phi_c)=\,V_0(\phi_c)-\sum_{\Gamma\in\,1\mathrm{PI}}\hbar^L\frac{U(\Gamma(p_1=0,\ldots,p_N=0))}{\sigma(\Gamma)}\frac{\phi_c^N}{N!} \tag{4.562}$$

where ϕ_c is viewed as a real variable, and $V_0(\phi_c)$ is the classical potential. By construction the quantum corrections are organized in increasing powers of $\hbar$ and these correspond to the loop number of the 1PI graphs. At the one-loop level and for a polynomial interaction, one finds that the unrenormalized value gives ([**267**], equation 2.64)

$$V(\phi_c) = V_0(\phi_c) + \frac{\hbar}{2}\int \log(1 + \frac{V_0''(\phi_c)}{k^2})\frac{d^D k}{(2\pi)^D} + O(\hbar^2) \tag{4.563}$$

In dimension $D = 4$ the integral diverges in the ultraviolet due to the two terms

$$\frac{V_0''(\phi_c)}{k^2} - \frac{(V_0''(\phi_c))^2}{2\,k^4} \tag{4.564}$$

in the expansion of the logarithm at large momentum k. If the classical potential V_0 is at most quartic the divergences can be compensated by adding suitable counterterms in the classical potential. Thus, in particular, if one uses an ultraviolet cutoff Λ and considers the ϕ^4 theory with classical potential $V_0(\phi) = \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}\phi^4$, one gets a quadratic divergence of the form

$$\frac{\Lambda^2}{32\pi^2}(3\lambda\phi_c^2 + m^2) - \frac{\log\Lambda}{32\pi^2}(V_0''(\phi_c))^2, \tag{4.565}$$

whose elimination requires adjusting the classical potential as a function of the cutoff Λ as

$$(V_0 + \delta V_0)(\phi) = V_0(\phi) - \frac{\Lambda^2}{32\pi^2}(3\lambda\phi^2) + \frac{\log\Lambda}{32\pi^2}(6m^2\lambda\phi^2 + 9\lambda^2\phi^4), \tag{4.566}$$

where we ignored an irrelevant (but Λ-dependent) additive constant.

This shows very clearly that, in order to obtain a Λ-independent effective potential, one needs the bare action to depend upon Λ with a large negative quadratic term of the form $-\frac{\Lambda^2}{32\pi^2}(3\lambda\phi^2)$ at the one-loop level. This is precisely the type of term present in the spectral action of Chapter I, in the case of the Standard Model. The presence of the other quadratic divergences coming from the Yukawa coupling of the scalar field with fermions alters the overall sign of the quadratic divergence only at small enough Λ. However, as shown in [**52**] §5.7, it reverts to the above sign when Λ gets above 10^{10} GeV and in particular when it is close to the unification scale.

After elimination of the divergences the renormalized effective potential at one-loop in the case of the Standard Model is of the form (cf. [**267**] equation 4.1)

$$\begin{aligned} V(\phi_c) = \; & \tfrac{1}{2}m^2\phi_c^2 + \tfrac{\lambda}{4}\phi_c^4 + \\ & \tfrac{\hbar}{64\pi^2}\Big[B\,\phi_c^4 \log\tfrac{\phi_c^2}{\mu^2} + (3\lambda\phi_c^2 + m^2)^2 \log\tfrac{(3\lambda\phi_c^2+m^2)}{\mu^2} \\ & +3(\lambda\phi_c^2 + m^2)^2 \log\tfrac{(\lambda\phi_c^2+m^2)}{\mu^2}\Big] + \dots \end{aligned} \tag{4.567}$$

where μ is the renormalization scale, and B is given by

$$B = \frac{3}{16}(3g_2^4 + 2g_1^2\,g_2^2 + g_1^4) - 3y_t^4. \tag{4.568}$$

One needs to check that the quantum correction does not affect too drastically the value $\phi_c = \sigma$ at which the first derivative of V vanishes. The classical value (assuming $m^2 < 0$) is given by $\sigma^2 = -\frac{m^2}{\lambda}$. The coefficient of the leading term $\phi_c^4 \log\frac{\phi_c^2}{\mu^2}$ in (4.567) is $X = \frac{\hbar}{64\pi^2}(12\lambda^2 + B)$. This term will give rise to an additional contribution to the derivative $dV/d\phi_c$ of the form $2X\phi_c^3$, which will add to the classical term $\lambda\phi_c^3$. This gives us as a rough estimate that, provided we take the running scale μ of the same order as σ, the leading alteration of σ^2 is obtained by replacing $\sigma^2 = -\frac{m^2}{\lambda}$ by $\sigma^2 = -\frac{m^2}{\lambda+2X}$ (cf. [**267**] equation 2.105). In the case of the Standard Model at the electroweak scale one has $2X \sim -\frac{1}{16\pi^2} \sim -0.0039$ while $\lambda \sim 0.24$, which gives an error of two percent.

8.1.2. *Effective potential at high temperature.* To understand how to take temperature into account, let us first give the physics interpretation of the one-loop correction to the effective potential in the zero temperature case. One considers (4.563) and performs the integration over the component k_0 of the Euclidean momentum. This gives (in $D = 4$ dimensions)

$$V(\phi_c) = V_0(\phi_c) + \hbar \int \sqrt{k^2 + V_0''(\phi_c)}\, \frac{d^3k}{(2\pi)^3} + O(\hbar^2), \tag{4.569}$$

which is the integral of the energy density of zero-point energy fluctuations around ϕ_c.

PROPOSITION 4.193. *The effective potential at temperature T in Euclidean spacetime is the same as the effective potential at zero temperature in the product of space by a circle S^1 of length $\beta = \frac{1}{T}$ with antiperiodic spin structure.*

PROOF. We repeat the formal proof given in [**267**, §3.1.2]. The theory at temperature T is given by the ensemble of temperature Green's functions associated to the Gibbs state

$$\begin{aligned} G = \;& \varphi_\beta(A_1(x_1,t_1)\cdots A_n(x_n,t_n)) \\ = \;& Z^{-1}\,\mathrm{Tr}(e^{-\beta H} A_1(x_1,t_1)\cdots A_n(x_n,t_n)) \end{aligned} \tag{4.570}$$

where $Z = \mathrm{Tr}(e^{-\beta H})$ and H is the Hamiltonian. Working at the formal level using the identity

$$A_m(x_m,t_m) = e^{it_m H} A_m(x_m,0) e^{-it_m H} \tag{4.571}$$

one obtains the following formula for the temperature Green's functions:

$$\begin{gathered} G((x_1,t_1),\ldots,(x_n,t_n)) = \\ Z^{-1}\,\mathrm{Tr}(e^{-\beta H} e^{it_1 H} A_1(x_1,0) e^{i(t_2-t_1)H} \cdots A_n(x_n,0) e^{-it_n H}) \end{gathered} \tag{4.572}$$

and one checks that this expression is invariant under the translation

$$it_m \mapsto it_m + \beta$$

In other words, the temperature Green's functions are periodic in Euclidean time with period β. Since the functional integral when performed in the product of space by a circle S^1 of length $\beta = \frac{1}{T}$ with antiperiodic spin structure does provide Green's functions which are periodic in Euclidean time with period β, one simply asserts that this gives the desired answer. □

Thus, in the computations of Feynman graphs in the perturbative expansion of the functional integral, one simply replaces integration over the k_0-component (dual to time) by

$$\int dk_0 \mapsto 2\pi T \sum_{n\in\mathbb{Z}} \quad \text{and} \quad \int dk_0 \mapsto 2\pi T \sum_{n\in\mathbb{Z}+\frac{1}{2}}, \tag{4.573}$$

according to whether one deals with bosons or fermions.

One can then compute the effective potential at temperature T at one loop. Instead of (4.563), it gives (cf. [**267**])

$$V(\phi_c) = V_0(\phi_c) + \frac{\hbar}{2}T\sum_{n\in\mathbb{Z}}\int \log(k^2+(2\pi nT)^2+V_0''(\phi_c))\frac{d^3k}{(2\pi)^3}+O(\hbar^2), \tag{4.574}$$

ignoring (infinite) constants such as $\frac{\hbar}{2}\int \log(k^2)\frac{d^4k}{(2\pi)^4}$. One then gets ([**267**])

$$V_T(\phi_c) = V_{T=0}(\phi_c) + \frac{\hbar T^4}{2\pi^2}\int_0^\infty \log\left(1-\exp[-\sqrt{x^2+V_0''(\phi_c)/T^2}\,]\right)x^2dx. \tag{4.575}$$

The right hand side is the free energy of a free field of bosons with mass $\frac{\sqrt{V_0''(\phi_c)}}{T}$, as one can see from Chapter I §2 (1.24). This holds for a scalar field and one has a similar expression for the contribution to the effective potential coming from the Yukawa coupling of the scalar field with fermions. This generates an additional term of the form

$$-4\frac{\hbar T^4}{2\pi^2}\int_0^\infty \log\left(1+\exp[-\sqrt{x^2+V_0''(\phi_c)/T^2}\,]\right)x^2dx, \tag{4.576}$$

which corresponds to the free energy of a free massive fermionic quantum field theory. One has a similar expression for the contribution to the effective potential coming from the minimal coupling with gauge bosons ([**267**]).

8.1.3. *Electroweak phase transition.* To analyze the high temperature behavior one expands the additional contributions to the effective potential above in decreasing powers of T. One obtains for these additional contributions at one loop, with the factor of $\hbar$ removed, the expressions

$$-\frac{11}{360}\pi^2T^4+\frac{1}{24}V_0''(\phi_c)T^2+\cdots \tag{4.577}$$

for bosons and

$$-\frac{7}{180}\pi^2T^4+\frac{1}{12}V_0''(\phi_c)T^2+\cdots \tag{4.578}$$

for fermions. (There is a typo in the corresponding expressions in [**267**].) Both expansions are valid for T^2 large with respect to $V_0''(\phi_c)$.

Thus, the effect of high temperature is to add a large positive term proportional to $T^2\phi_c^2$ to the effective potential, and for high enough T this will eliminate the symmetry breaking vacuum. We then get a situation which is reminiscent of the phase transition occurring in the quantum statistical mechanical systems of Chapter III, with symmetry being restored at high enough temperature. In order to get a better understanding of a possible extension of this phenomenon to the full gravitational sector (of which the electroweak sector is only a small "inner" part) we need to describe in a more precise manner the general conceptual framework for QG, in analogy with the framework of Chapter III.

8.2. Observables in QG.

We saw in Chapter I §2 that the classification of vacuum states on the algebra of observables in a given QFT provides a natural framework for the non-perturbative formulation of the theory. Moreover, the classification of KMS_β states is the natural extension of the theory at finite temperature. This can be seen in the details of the proof of Proposition 4.193 and it provides at the formal level the right algebraic framework to formulate the symmetry breaking associated to the analysis of the effective potential in §8.1. One difficulty in this approach is to obtain a well-defined time evolution, since the Hamiltonian is a priori only defined in a perturbative manner.

Table 2 suggests an analogy between the well-defined quantum statistical mechanics of $\mathbb{Q}$-lattices analyzed in Chapter III and an extension of the above algebraic framework of QFT to the full gravitational sector. We need to keep in mind that what plays the role of $\mathbb{Q}$-lattices in this analogy is space, and not space-time. Thus we are discussing here the geometry of 3-dimensional space rather than 4-dimensional space-time. In particular, the spectral observables such as $D \mapsto \mathrm{Tr}(D^{-n}) = \mathrm{Tr}(ds^n)$, which play the same role as the Eisenstein series of Chapter III, are invariants of the Dirac Hamiltonian rather than of the Euclidean fermion propagator.

The idea here is to develop a suitable notion of "spectral correspondences" Γ that generalizes spectral triples but relates two different geometries, i.e. involves two different algebras. We elaborate on this idea of spectral correspondence in §8.4 below. In fact it is still unclear what the relevant notion of spectral correspondence should be and it could well be formulated in terms of 4-dimensional cobordisms between 3-geometries as explained in §8.5.

The composition of spectral correspondences Γ should allow one to define a convolution algebra $\mathcal{H}_G$ of functions $f(\Gamma)$ similar to the convolution algebra of the groupoid of commensurable pairs of $\mathbb{Q}$-lattices, up to scale. This Hecke algebra $\mathcal{H}_G$ should be closely related to the algebra of observables in the still non-existing theory of quantum gravity.

8.3. Invertibility at low temperature.

We have seen in Chapter 3 how the invertible $\mathbb{Q}$-lattices, seen as the classical points of the Shimura variety of GL_1 or GL_2 appear as extremal KMS states at low temperature for the algebra of coordinates on the noncommutative space of $\mathbb{Q}$-lattices up to commensurability and scaling with its natural time evolution defined by the ratio of covolumes of commensurable lattices. The analogous idea here is that a spectral geometry, such as the one we used in Chapter I to describe the Standard Model coupled to gravity at unification scale, would only appear as a low temperature state of an algebra of coordinates on the space of all spectral correspondences up to suitable equivalence. The phase transition phenomena we observed, in Chapter III, in the case of $\mathbb{Q}$-lattices suggest that, at a sufficiently high temperature, one can have an unbroken phase (or even one may not have KMS states at all as in the case of the GL_2 system).

This, in this context of quantum gravity, would mean that geometry may only appear as an emergent phenomenon coming from a low temperature state. In particular, what this analogy suggests is that, when applying the formalism of quantum mechanics to gravitation, i.e. to the entire universe, one should not use ground states

$\mathbb{Q}$-lattices	**Quantum Gravity**
$\mathbb{Q}$-lattice	Real spectral triple of dimension $d = 3$
Invertible $\mathbb{Q}$-lattice	Poincaré duality spectral triple
Commensurable pair of $\mathbb{Q}$-lattices	Spectral correspondence (or cobordism)
Scaling action	$D \mapsto \lambda D$
Composition of pairs	Composition of correspondences
$\gamma \mapsto \gamma^{-1}$	Contragredient correspondence
Groupoid C^*-algebra	Hecke algebra of functions of correspondences
Eisenstein series	$D \mapsto \mathrm{Tr}(D^{-n}) = \mathrm{Tr}(ds^n)$
Shimura variety	Moduli space of Dirac operators

TABLE 2. $\mathbb{Q}$-lattices versus Quantum Gravity

but rather temperature states. It raises in particular the possibility that the geometry of space only emerges after a suitable symmetry breaking mechanism which extends to the full gravitational sector the electroweak symmetry breaking that we discussed above. The invariance of the spectral action under the symplectic unitary group in Hilbert space is broken during this process to the compact group of isometries of a given geometry. In order to have an unbroken phase it is important to include among configurations those in which the line element ds can be degenerate and in particular non-invertible including the extreme case $ds = 0$. This is in line with the need to include non-invertible vierbeins when dealing with $2+1$ gravity ([**301**], [**302**]). What we saw in Chapter III, in the context of $\mathbb{Q}$-lattices, is that the invertible $\mathbb{Q}$-lattices have positive measure for the extremal KMS_β states only at low enough temperature. In the context of geometry, among the conditions fulfilled by the Dirac spectral triple of ordinary smooth Riemannian manifolds, as formulated in [**70**] and recalled in Chapter I §18.2, the following two involve an "invertibility":

- The real orientability (cf. Chapter I §18.1).
- The invertibility of the intersection form (cf. Chapter I §18.2).

These are *invertibility* conditions which ought to play a role similar to the invertibility for $\mathbb{Q}$-lattices. Thus, they should not be assumed in general but should single out special elements from which positive energy representations can be constructed.

8.4. Spectral correspondences.

We first briefly explain how the notion of real spectral triple (Chapter I) extends to the situation in which instead of a bimodule over the same algebra $\mathcal{A}$ one considers the asymmetric case of two different algebras $\mathcal{A}, \mathcal{B}$. In fact it suffices to consider the algebra $\mathcal{C} = \mathcal{A} \oplus \mathcal{B}$ and work with real spectral triples for $\mathcal{C}$ for which the representation of $\mathcal{C} \otimes \mathcal{C}^0$ is supported by the projection $\mathbf{1}_\mathcal{A} \otimes \mathbf{1}^0_\mathcal{B} + \mathbf{1}_\mathcal{B} \otimes \mathbf{1}^0_\mathcal{A}$ where $\mathbf{1}_\mathcal{A}$ and $\mathbf{1}_\mathcal{B} = 1 - \mathbf{1}_\mathcal{A}$ are the central projections. One can then isolate the structure corresponding to the subspace which is the range of $\mathbf{1}_\mathcal{A} \otimes \mathbf{1}^0_\mathcal{B}$. It is by construction an $\mathcal{A} - \mathcal{B}$-bimodule. Passing to the contragredient bimodule (cf. Definition 1.174 of Chapter I) should allow one to obtain a natural involution on the algebra $\mathcal{H}_G$. It also should give a natural reason for the real structure J, which ought to be viewed as an isomorphism of the correspondence with its contragredient. This is the counterpart of the condition $\gamma = \gamma^{-1}$ which is fulfilled by units in a groupoid.

REMARK 4.194. The above notion of spectral correspondence should play a role in extending to the noncommutative case the notion of fundamental group. Indeed one expects that this analogue should be incarnated as a suitable tensor category (like the category of finite dimensional representations of a group). One thus expects to deal with suitable $\mathcal{A}$–$\mathcal{B}$-bimodules which fulfill a suitable "flatness" condition.

8.5. Spectral cobordisms.

Rather than using the idea of correspondence described above in §8.4, there is an alternative which is to relate two 3-geometries by cobordisms using 4-geometries *with boundary.* The case of geometries with boundary is developed in [**51**] where it is shown that manifolds with boundary naturally give rise, in the even dimensional case, to a spectral triple. Moreover the spectral action delivers the York–Gibbons–Hawking boundary term with correct sign and relative coefficient. This suggests to develop, in the context of spectral triples, an algebra of cobordisms between 3-geometries, where a cobordism is weighted by the coefficient $e^{-S(D/\Lambda)}$, using the spectral action. The resulting algebra then requires performing a sum which is similar to the Euclidean functional integral.

8.6. Scaling action.

If we follow the analogy with the case of $\mathbb{Q}$-lattices, we need to divide by the scaling. Note that there is a natural action of $\mathbb{R}_+^*$ on spectral triples which is simply given by

$$(\mathcal{A}, \mathcal{H}, D) \mapsto (\mathcal{A}, \mathcal{H}, \lambda D), \quad \forall \lambda \in \mathbb{R}_+^* . \tag{4.579}$$

This action should extend to spectral correspondences (and to cobordisms) exactly as the scaling action on $\mathbb{Q}$-lattices extends to commensurable pairs. Moding out by the scaling action is tantamount to focussing on the conformal structure underlying the geometry.
For each $s \in \mathbb{C}$ one has a function which is homogeneous of weight s,

$$E_s(\mathcal{H}, D) = \mathrm{Tr}(|D|^{-s}) \tag{4.580}$$

which is the analogue of the function on $\mathbb{Q}$-lattices given by

$$\sum_L |y|^{-s}$$

The analogue of the Eisenstein series is

$$E(n)(\mathcal{H}, D) = \mathrm{Tr}(D^{-n}). \tag{4.581}$$

We need to understand the meaning of the dual system, as explained in §4.9 of Chapter III. Of course we expect the dual action to be related to the renormalization group flow $\mu\partial_\mu$.
Having a good notion of dual system and scaling action makes all the tools of cooling and distillation described earlier in this chapter available to this context, including the notion of classical points of the periodic orbits that we discussed in the case of the Bost–Connes system in relation to the spectral realization of zeros of the Riemann zeta function. It is still unclear at this stage how these notions should be interpreted in the context of quantum gravity, but one can follow the example we developed in studying the Riemann zeta function as a guideline. In Chapter IV §4, we have seen how the spectral realization is obtained naturally using the distillation procedure from the dual system. It is at this level of the construction that one expects to make contact with the BRS cohomological techniques which are used in QFT to deal with situations involving gauge invariance.

8.7. Moduli spaces for $\mathbb{Q}$-lattices and spectral correspondences.

Another interesting comparison between the case of the quantum statistical mechanical systems for $\mathbb{Q}$-lattices and the proposed approach to quantum gravity is that similar moduli spaces may play a role in both cases. As we have seen in Chapter 3 (cf. also the work of [**159**]), for the quantum statistical mechanical systems obtained by considering the equivalence relation of commensurability on $\mathbb{Q}$-lattices, the extremal low temperature KMS states are parameterized by a classical moduli space which has the form of an adelic quotient. In the case of 1-dimensional and 2-dimensional $\mathbb{Q}$-lattices this is given, respectively, by the Shimura varieties $\mathrm{GL}_1(\mathbb{Q})\backslash(\mathrm{GL}_1(\mathbb{A}_{\mathbb{Q},f}) \times \{\pm 1\})$ and $\mathrm{GL}_2(\mathbb{Q})\backslash(\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q},f}) \times \mathbb{H}^{\pm}) = \mathrm{GL}_2(\mathbb{Q})\backslash \mathrm{GL}_2(\mathbb{A}_Q)/\mathbb{C}^*$. On the other hand, as we have seen in Chapter 1, the moduli space of Dirac operators on the finite geometries associated to the matter sector of the Standard Model coupled to gravity is given by a quotient of a somewhat similar nature, namely by the double quotient $(U(3) \times U(3)) \backslash (\mathrm{GL}_3(\mathbb{C}) \times \mathrm{GL}_3(\mathbb{C}))/U(3)$ for the quark sector and a fibration over it for the lepton sector (cf. Chapter 1, Proposition 1.192). In addition to this kind of moduli space for the matter sector, one should expect to have in the algebra of spectral correspondences also the 3-manifolds that give the gravitational part (as we mentioned above, we think here of working with space rather than spacetime). One can imagine, for instance, that the units of the groupoids may be taken to be product geometries of S^3 by a finite geometry of the type described in Chapter 1, while the arrows would be be spectral correspondences that are again products of a finite spectral correspondence by a 3-manifold, thought of as a geometric correspondence over the 3-sphere, for instance by realizing it as a covering of the 3-sphere branched along a knot. It may be that one has to extend the notion of spectral triple and spectral correspondence to incorporate the boundary strata in a suitable compactification of the moduli space of Dirac operators of Chapter 1, Proposition 1.192, and possibly adelic versions as in the $\mathbb{Q}$-lattices setting.

Appendix

This appendix covers some basic notions and material that we need in the rest of the book.
We provide a quick reminder of some basic notions of operator algebras that we use as tools for the analytic description of noncommutative spaces that we deal with throughout the book. We also recall some notions from number theory, especially the general ideas of Galois theory, which is a recurrent theme we encounter in different contexts and settings.

9. Operator algebras

9.1. C^*-algebras.

The use of C^*-algebras to describe noncommutative spaces originates with the work of Israel Gelfand. In fact, noncommutative C^*-algebras provide a natural extension of the algebro–geometric duality

$$\text{Space} \leftrightarrow \text{Commutative algebra.} \tag{4.582}$$

This duality can be formulated in the form of the Gelfand–Naimark correspondence

$$X \leftrightarrow C_0(X)$$

between locally compact Hausdorff topological spaces and commutative C^*-algebras.

Definition 4.195. *An involutive algebra $\mathcal{A}$ is an algebra over $\mathbb{C}$, together with an involution $* : \mathcal{A} \to \mathcal{A}$ satisfying*

$$(x^*)^* = x, \quad (x+y)^* = x^* + y^*, \quad (xy)^* = y^* x^*, \quad (\lambda x)^* = \bar{\lambda} x^*, \tag{4.583}$$

for all $x, y \in \mathcal{A}$ and all $\lambda \in \mathbb{C}$.

Definition 4.196. *A C^*-algebra $\mathcal{A}$ is an involutive algebra which is complete with respect to a Banach algebra norm $\|\cdot\|$ such that the C^*-condition holds, namely*

$$\|xx^*\| = \|x\|^2 \quad \forall x \in \mathcal{A}\,. \tag{4.584}$$

Note that the norm $\|\cdot\|$ is uniquely determined by the algebraic structure and is in fact given by the spectral radius:

$$\|x\| = \sup\{\lambda \in \mathbb{R}^+ \ : \ xx^* - \lambda^2 \notin \mathcal{A}^{-1}\}\,. \tag{4.585}$$

where $\mathcal{A}^{-1}$ denotes the set of invertible elements of $\mathcal{A}$.
The Gelfand–Naimark correspondence for commutative C^*-algebras assigns to $\mathcal{A}$ the set X of its maximal ideals. This is a locally compact Hausdorff space, compact when the algebra is unital. By identifying X with the set of homomorphisms $\phi : \mathcal{A} \to \mathbb{C}$, the Gelfand transform $a \mapsto f_a(\phi) = \phi(a)$ gives an isometric $*$-isomorphism $\mathcal{A} \cong C_0(X)$, so that every commutative C^*-algebra has this form.

An important invariant of C^*-algebras is given by the K-groups $K_0(\mathcal{A})$ and $K_1(\mathcal{A})$. The $K_0(\mathcal{A})$ group is the algebraic K_0 theory of the ring $\mathcal{A}$, namely it is the Grothendieck group of the semigroup $K_0(\mathcal{A})_+$ generated by projections in matrix algebras $M_n(\mathcal{A})$ up to stable equivalence, or equivalently stable isomorphism classes of finite projective modules over $\mathcal{A}$. The group $K_1(\mathcal{A})$ is obtained by considering the invertible elements $\mathrm{GL}_n(\mathcal{A})$ in $M_n(\mathcal{A})$, modulo the connected component $\mathrm{GL}_n(\mathcal{A})^0$. The direct limit of the induced homomorphisms $\mathrm{GL}_n(\mathcal{A})/\mathrm{GL}_n(\mathcal{A})^0 \to \mathrm{GL}_{n+1}(\mathcal{A})/\mathrm{GL}_{n+1}(\mathcal{A})^0$ defines $K_1(\mathcal{A})$. For a separable C^*-algebra, K_0 and K_1 are countable abelian groups. In the case of a commutative algebra $C(X)$, they give the topological K-theory groups $K^i(X)$.

9.2. Von Neumann algebras.

One should think of von Neumann algebras as algebras of operators *defined by symmetries*. This means that a von Neumann algebra $\mathcal{M}$ is a unital involutive algebra of operators with the property that there is a group $\mathcal{U}$ of unitary transformations such that the elements $T \in \mathcal{M}$ are characterized by the invariance $UTU^* = T$ for all $U \in \mathcal{U}$:

$$\mathcal{M} = \{T \,|\, UTU^* = T\,, \quad \forall U \in \mathcal{U}\} \tag{4.586}$$

More precisely, let $\mathcal{H}$ be a separable Hilbert space and let $\mathcal{L}(\mathcal{H})$ denote the algebra of bounded linear operators on $\mathcal{H}$.

DEFINITION 4.197. *The* commutant *of a subset $\mathcal{M} \subset \mathcal{L}(\mathcal{H})$ is defined as*

$$\mathcal{M}' := \{T \in \mathcal{L}(\mathcal{H}) : TS = ST, \ \forall S \in \mathcal{M}\}. \tag{4.587}$$

A von Neumann algebra is then defined as follows.

DEFINITION 4.198. *A von Neumann algebra is a unital subalgebra $\mathcal{M} \subset \mathcal{L}(\mathcal{H})$, which is stable under adjoint $T \mapsto T^*$ and satisfies one of the following equivalent conditions.*

(1) *$\mathcal{M} = \mathcal{M}''$, where $\mathcal{M}'' = (\mathcal{M}')'$ is the double commutant.*
(2) *$\mathcal{M}$ is closed in the weak topology.*
(3) *$\mathcal{M}$ is closed in the strong topology.*

A net (T_α) of operators $T_\alpha \in \mathcal{L}(\mathcal{H})$ is convergent to T for the weak topology (resp. the strong topology) if and only if $\langle T_\alpha \xi, \eta\rangle \to \langle T \xi, \eta\rangle$ (resp $T_\alpha \xi \to T\xi$) for any vectors $\xi, \eta \in \mathcal{H}$. The σ–weak (and σ–strong) topology is defined in a similar manner using absolutely convergent countable sums, i.e. $\sum \langle T \xi_n, \eta_n\rangle$ instead of $\langle T \xi, \eta\rangle$.
The equivalence of the three conditions in the above definition is the content of von Neumann's double commutant theorem. This equivalence holds for any of the four possible topologies: weak, strong, σ–weak and σ–strong. In particular, the commutant provides a way of constructing von Neumann algebras. For the pair of von Neumann algebras $\mathcal{M}$ and $\mathcal{M}'$, given by $\mathcal{M}$ and its commutant, the intersection

$$\mathcal{M} \cap \mathcal{M}' = Z(\mathcal{M}) = Z(\mathcal{M}')$$

is the center for both algebras.
Finally, von Neumann algebras can be described as dual spaces, namely, we have the following.

PROPOSITION 4.199. *Let $\mathcal{L}^1(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$ denote the trace class operators on $\mathcal{H}$. Let $\mathcal{M} \subset \mathcal{L}(\mathcal{H})$ be a von Neumann algebra, and $\mathcal{M}^\perp$ be the annihilator*

$$\mathcal{M}^\perp = \{T \in \mathcal{L}^1(\mathcal{H}) : \operatorname{Tr}(ST) = 0, \forall S \in \mathcal{M}\}.$$

Then the Banach space $\mathcal{M}_* = \mathcal{L}^1(\mathcal{H})/\mathcal{M}^\perp$ is the predual of $\mathcal{M}$, that is, $(\mathcal{M}_*)^* = \mathcal{M}$. Under this identification, the σ–weak topology on $\mathcal{M}$ is the one induced by the functionals $S \mapsto \operatorname{Tr}(ST)$, for some $T \in \mathcal{M}_*$, and for all $S \in \mathcal{M}$.

In fact, one can show that, for any unital subalgebra $\mathcal{M} \subset \mathcal{L}(\mathcal{H})$ stable under adjoint, the double commutant $\mathcal{M}''$ is the σ–weak closure of $\mathcal{M}$. This, in particular, implies that von Neumann algebras are C^*-subalgebras of $\mathcal{L}(\mathcal{H})$ that are σ–weak closed. It is also possible to characterize von Neumann algebras as precisely those C^*-algebras that are dual spaces.
One should keep in mind that, though this in general implies that it is possible to regard von Neumann algebras as C^*-algebras, this always (except in the finite dimensional case) gives rise to C^*-algebras that are non–separable, hence not really relevant (except possibly as intermediate tools) as noncommutative topological spaces (somewhat like non–metrizable topological spaces). To maintain the correct intuition, one should rather think of von Neumann algebras as measure spaces, and separable C^*-algebras as metrizable topological spaces.

In von Neumann's setting, ordinary measure theory becomes the special case of commutative von Neumann algebras, in the following sense.

EXAMPLE 4.200. *Let (X, μ) be a measure space, with X a locally compact Hausdorff space, and μ a probability measure on the Borel sets of X. We can regard functions f in $L^\infty(X, \mu)$ as multiplication operators M_f acting as bounded operators on $L^2(X, \mu)$. This gives rise to a commutative von Neumann algebra $\mathcal{M} = L^\infty(X, \mu)$.*

In fact, this example describes the general commutative von Neumann algebras up to isomorphism. Measure theory can be reconstructed from commutative von Neumann algebras. This is a consequence of the spectral theorem and can be seen as follows.

DEFINITION 4.201. *Let $\mathcal{M} \subset \mathcal{L}(\mathcal{H})$ be a unital subalgebra closed under adjoint. A vector $x \in \mathcal{H}$ is cyclic for $\mathcal{M}$ if the subspace $\mathcal{M}x$ is dense in $\mathcal{H}$. A vector $x \in \mathcal{H}$ is separating for $\mathcal{M}$ if it is cyclic for the commutant $\mathcal{M}'$.*

We then have the following result.

PROPOSITION 4.202. *Consider an operator T in $\mathcal{L}(\mathcal{H})$, which is normal (commuting with its adjoint), and consider the C^*-algebra $C^*(T)$ generated by T. The minimal von Neumann algebra $W^*(T)$ containing T and the identity I will then be the closure of $C^*(T)$ in the strong topology. Assume $W^*(T)$ has a cyclic vector. Then there is an isometry U of $L^2(X, \mu)$ with $\mathcal{H}$, where $X = \sigma(T)$ is the spectrum of T and μ is the spectral measure, and $f \mapsto UM_fU^*$ gives an isometric $*$–isomorphism between $L^\infty(X, \mu)$ and $\mathcal{M}$.*

This isomorphism is obtained by assigning to each polynomial

$$p(z, \bar{z}) = \sum a_{ij} z^i \bar{z}^j$$

the operator $p(T, T^*) = \sum a_{ij} T^i (T^*)^j$ in the algebra $W^*(T)$. This correspondence extends to all Borel functions f, and one shows that for a bounded sequence of functions f_n which converges pointwise to f on $X = \sigma(T)$, the operators $f_n(T)$ converge strongly to an operator which defines $f(T)$.

We say that an operator S has the same symmetries as T if it satisfies $USU^* = S$ for any unitary operator U on $\mathcal{H}$ such that $UTU^* = T$. When T is normal any such operator S can be written in the form $S = f(T)$, for some Borel function f.

The condition on the existence of a cyclic vector for $\mathcal{M}$ corresponds to being maximal commutative, i.e. not contained in any larger commutative $*$–subalgebra of $\mathcal{L}(\mathcal{H})$. In turn this is equivalent to the condition $\mathcal{M} = \mathcal{M}'$, and corresponds in the spectral theorem to requiring eigenvalues to be simple. In a more general case, one needs to consider the additional datum of a multiplicity function at points of $X = \sigma(T)$.

We have seen that commutative von Neumann algebras recover classical measure theory. On the other hand, one can consider von Neumann algebras that are "as far as possible" from the commutative case. This gives rise to the very rich theory of *factors*.

For a general von Neumann algebra $\mathcal{M}$ the center $Z(\mathcal{M})$ is always of the form $L^\infty(X, \mu)$ for some measure space (X, μ). The case which is as noncommutative as possible is when this center consists only of constant functions.

DEFINITION 4.203. *A von Neumann algebra $\mathcal{M}$ with trivial center, $Z(\mathcal{M}) = \mathbb{C}$, is called a* factor.

An important notion in studying factors is that of a *trace*.

DEFINITION 4.204. *Let $\mathcal{M}$ be a factor. A trace τ on $\mathcal{M}$ is a linear functional on $\mathcal{M}$ satisfying $\tau(ab) = \tau(ba)$ for all $a, b \in \mathcal{M}$. A trace is normalized if it satisfies the further condition $\tau(1) = 1$.*

The term factor was introduced by von Neumann, whose aim in relation with quantum mechanical systems was to show that such an algebra should induce a factorization $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ of the ambient Hilbert space $\mathcal{H}$ such that $\mathcal{M}$ consists of the operators of the form $T_1 \otimes 1$, with $T_1 \in \mathcal{L}(\mathcal{H}_1)$. As shown early on by Murray and von Neumann such a factorization result does not hold in general. An example where it fails can be given as follows. Consider a discrete group Γ, with the property that the conjugacy classes of all non–trivial elements in Γ are infinite. Consider the Hilbert space $\ell^2(\Gamma)$ with canonical basis $(\epsilon_h)_{h\in\Gamma}$ and the left regular representation

$$\lambda(g)\epsilon_h = \epsilon_{gh}.$$

The algebra $\mathcal{M}$ of left-invariant operators on $\ell^2(\Gamma)$ is a factor. Moreover, the function

$$T \in \mathcal{M} \mapsto \langle \epsilon_1, T\epsilon_1 \rangle =: \tau(T)$$

defines a finite trace, with $\tau(I) = 1$. However, if $\mathcal{M}$ were the algebra of bounded operators on a Hilbert space, there would be no such finite trace.

Though the factorization result originally expected by von Neumann does not hold, factors are indeed crucial to the study of von Neumann algebra, because all information about von Neumann algebras can be reduced to the case of factors, via von Neumann's reduction theorem. In fact, in [**288**], von Neumann proved the following:

THEOREM 4.205. *A von Neumann algebra $\mathcal{M}$ on a separable Hilbert space $\mathcal{H}$ is always isomorphic to a direct integral of the form*

$$\int_X \mathcal{M}(x) d\mu(x)$$

over a measure space (X, μ), where each $\mathcal{M}(x)$ is a factor.

Murray and von Neumann classified factors into the types I_n and I_∞, II_1 and II_∞, and III. Their initial classification produced the following picture.

Type I_n	$M_n(\mathbb{C})$
Type I_∞	$\mathcal{L}(\mathcal{H})$
Type II_1	factors that admit a finite trace
Type II_∞	$\mathrm{II}_1 \otimes \mathrm{I}_\infty$
Type III	whatever remains

In this picture, factors of Type I are very rigid: either finite dimensional matrix algebras, or the algebra of all bounded operators on a Hilbert space. There is a lot more interesting structure in Type II and III factors. For a long time, however, very few examples were known, until Powers introduced a continuous family of mutually non–isomorphic type III factors [**246**]. It was later understood that the best way to approach the theory of factors and produce many significant examples is geometrically, via the theory of foliations [**66**]. All types of factors occur from this geometric construction and the continuous dimensions of Murray and von Neumann play an essential role in the longitudinal index theorem [**98**]. Typical examples of algebras of foliations that are of Type I occur when the foliation is actually a fibration, or when all the leaves are compact, or in slightly more general cases like the Reeb foliation of the 3-sphere, where in any case the foliation admits a measurable section and the space of leaves is a nice Borel space. Examples of Type II include the Kronecker foliation that defines noncommutative tori. An example of a foliation that gives a Type III case is the Anosov foliation given by the orbits of the right action of upper triangular matrices on the unit tangent bundle $\Gamma\backslash\mathrm{PSL}(2,\mathbb{R})$ of a compact Riemann surface of genus $g > 1$ with Fuchsian uniformization $\Sigma = \Gamma\backslash\mathbb{H}$.

While the existence of traces provides a tool for the study of type II factors, it was only the use of Tomita theory that made an in-depth study and a classification of type III factors possible [**63**]. We will not discuss in detail the classification of factors here and we refer the interested reader to [**63**]. However, we recall an important result of the theory of von Neumann algebras, which is used several times in the rest of this book and we show how it came to play an important role in the classification of factors.

9.3. The passing of time.

An important property of noncommutative von Neumann algebra is that they evolve in time, namely, they carry a one-parameter group of automorphisms, which is canonical up to inner automorphisms [**63**].

Recall that, in classical and in quantum mechanics, the notion of time can be introduced both as a flow on the configuration space or as a one-parameter group of automorphisms of the algebra of observables. In quantum mechanics, these two approaches represent the complementary viewpoints of Schrödinger and Heisenberg.

On the C^*-algebras of observables $\mathcal{A}$ of a quantum mechanical system, represented as an algebra of operators on a Hilbert space $\mathcal{H}$, a positive trace-class operator ω on

$\mathcal{H}$ determines a state of the form

$$\varphi_\omega(a) = \frac{\mathrm{Tr}(a\omega)}{\mathrm{Tr}(\omega)} \quad \forall a \in \mathcal{A}. \tag{4.588}$$

A von Neumann algebra $\mathcal{M}$ is then obtained by considering the closure of $\mathcal{A}$ in the topology determined by this trace-class operator.
Let ξ be a cyclic and separating vector in $\mathcal{H}$, so that we can write the state as a quantum mechanical expectation

$$\varphi_\omega(a) = \langle \xi, a\xi \rangle. \tag{4.589}$$

The (closure of) the operator defined by

$$S a\, \xi = a^*\xi, \tag{4.590}$$

for all $a \in \mathcal{M}$, admits a polar decomposition of the form

$$S = J\Delta^{1/2}, \tag{4.591}$$

with Δ a positive self-adjoint operator and J anti-unitary. Tomita's theory [**279**] provides the following result.

THEOREM 4.206. (Tomita) *The operator S of* (4.590) *has the following properties.*

$$\begin{aligned} J\mathcal{M}J &= \mathcal{M}', \\ \Delta^{-it}\mathcal{M}\Delta^{it} &= \mathcal{M}. \end{aligned} \tag{4.592}$$

The first part of this theorem should be read as the statement that "$\mathcal{M}$ has the same size of its commutant". The second part provides a one–parameter group of automorphisms σ_t^φ of the von Neumann algebra $\mathcal{M}$, given a state φ, hence the notion of a time flow in the sense of the Heisenberg picture. When we introduce the formalism of quantum statistical mechanics, we also introduce an important class of states that correspond to the possible thermodynamical equilibrium states of the system at a given inverse temperature β. These are characterized by the KMS (Kubo–Martin–Schwinger) condition, which extends the notion of Gibbs measure for classical statistical mechanical systems. We then return to the time evolution σ_t^φ defined by Theorem 4.206. Takesaki [**279**] showed that it is the unique one-parameter group of automorphisms of $\mathcal{M}$ that satisfies the KMS condition for $\beta = 1$. An even more surprising fact is true. Connes proved in [**63**] that, modulo the inner automorphisms of the algebra, the one-parameter group σ_t^φ determines a *canonical time evolution*, which no longer depends on the choice of the state φ. The normal subgroup of inner automorphisms $\mathrm{Inn}(\mathcal{M})$ is given by automorphisms realized by unitary operators,

$$T \mapsto U^*TU.$$

The main result on the Tomita time evolution σ_t^φ is given by the following theorem.

THEOREM 4.207. (Connes [**63**]) *In the exact sequence*

$$1 \to \mathrm{Inn}(\mathcal{M}) \to \mathrm{Aut}(\mathcal{M}) \to \mathrm{Out}(\mathcal{M}) \to 1,$$

the class of σ_t^φ in $\mathrm{Out}(\mathcal{M})$ *does not depend upon the choice of the state φ.*

Thus, given a von Neumann algebra $\mathcal{M}$, there is a completely canonical homomorphism

$$\mathbb{R} \xrightarrow{\delta} \mathrm{Out}(\mathcal{M}). \tag{4.593}$$

This fact is entirely an effect of noncommutativity, with no classical analog for commutative algebras.

In [**97**] a hermeneutical interpretation for the time flow (4.593) is proposed as the physical time in generally covariant theories, and this hypothesis is tested against non–relativistic and classical limits, and cosmology. In fact, a major problem arises in trying to recover the notion of physical time in generally covariant theories, such as general relativity, and any theory involving the gravitational field relativistically, where no preferred time flow can be assigned. In [**97**] the idea of [**256**] of a definition of physical time that depends on the thermal state of the system is given a rigorous mathematical content in the form of the Tomita time evolution of von Neumann algebras.

This independence of the state of the automorphism group satisfying the KMS condition modulo inner automorphisms plays a crucial role in Connes' classification of Type III factors in [**63**].

In fact, it is possible to define invariants of factors using the group of modular automorphisms

$$\sigma_t^\varphi(a) = \Delta_\varphi^{-it} a \Delta_\varphi^{it} \tag{4.594}$$

associated to the state $\varphi(a) = \langle x, ax \rangle$. Two important invariants are defined by

$$S(\mathcal{M}) := \cap_\varphi \mathrm{Spec}(\Delta_\varphi) \tag{4.595}$$

$$T(\mathcal{M}) := \{t \in R \,:\, \sigma_t^\varphi \in \mathrm{Inn}(\mathcal{M})\}. \tag{4.596}$$

In principle, it would be extremely hard to compute invariants such as $S(\mathcal{M})$, since they involve all states. However, the result of Theorem 4.207 provides the key to make such invariants computable.

The multiplicative subgroup $S(\mathcal{M}) \setminus \{0\}$ of $\mathbb{R}_+^*$ subdivides type III factors into the types III_λ for which $S(\mathcal{M}) = \{\lambda^n \,:\, n \in \mathbb{Z}\} \cup \{0\}$, with $0 < \lambda < 1$, or $S(\mathcal{M}) = \mathbb{R}_+^*$ for $\lambda = 1$, or $S(\mathcal{M}) = \{0, 1\}$, for $\lambda = 0$.

Type III factors can be studied by reducing them to the type II case via a crossed product construction involving the time evolution. This comes up crucially in this book when we discuss quantum statistical mechanical systems like the Bost–Connes system with arithmetic properties and the dual system obtained by taking the crossed product by the time evolution. The underlying structure theory is exactly the reduction from type III factors to type II factors and their automorphisms first introduced by Connes in [**63**].

The classification result for factors of type III_λ with $0 < \lambda < 1$ is then of the following form.

Theorem 4.208. (Connes, [**63**]) *For $0 < \lambda < 1$, type III_λ factors have a discrete decomposition of the form $\mathcal{M} = \mathcal{N} \rtimes_\theta \mathbb{Z}$, with $\mathcal{N}$ a type II_∞ factor, and θ an automorphism of $\mathcal{N}$, acting on any trace τ on $\mathcal{N}$ by scaling $\tau \circ \theta = \lambda \tau$. Conversely, every such crossed product gives a type III_λ factor. The pair $(\mathcal{N}, \theta)$ is uniquely determined up to inner automorphisms.*

In the Type III case, the role of traces is replaced by generalized traces, namely semi-finite normal weights φ such that $\varphi(1) = \infty$ and $\mathrm{Spec}(\Delta_\varphi) = S(\mathcal{M})$. The factor $\mathcal{N}$ is the centralizer of a generalized trace.

By theorem 4.208 the classification of type III_λ factors is reduced to classification of type II_∞ factors and of classes of θ in $\mathrm{Out}(\mathcal{N})$ with the property that $\tau \circ \theta = \lambda\tau$. This can be treated within the framework of noncommutative ergodic theory (cf. [**68**] for a detailed discussion of these topics).

In the type III_0 case, the result is similar to Theorem 4.208 above, but the algebra $\mathcal{N}$ need no longer be a factor and the uniqueness result is then given in terms of conjugate automorphisms induced by projections in the center.

THEOREM 4.209. (Connes [**63**]) *Type III_0 factors have a discrete decomposition of the form $\mathcal{M} = \mathcal{N} \rtimes_\theta \mathbb{Z}$, with $\mathcal{N}$ of type II_∞ (not necessarily a factor), and θ an automorphism of $\mathcal{N}$, which is a contraction (i.e. $\tau \circ \theta \leq \lambda\tau$ with $\lambda < 1$). Given two such pairs $(\mathcal{N}_j, \theta_j)$, $j = 1, 2$, which give rise to the same $\mathcal{M}$, there are projections e_j in $Z(\mathcal{N})$ such that the induced automorphisms θ_{j,e_j} are conjugate.*

The case of type III_1 can be derived from the continuous decomposition of type III von Neumann algebras ([**280**]).

10. Galois theory

As shown repeatedly throughout this book, many interactions between noncommutative geometry and number theory involve some form of Galois theory. Instances include *motivic* Galois theory and *differential* Galois theory that play a crucial role in the theory of perturbative renormalization in quantum field theory. Another important form of Galois theory that is interestingly related to noncommutative geometry is *class field theory* and Hilbert's 12th problem (explicit class field theory). In particular we are using the Kronecker–Weber theory of the maximal abelian extension of $\mathbb{Q}$, i.e. the Galois theory of cyclotomic fields, as well as the explicit class field theory for imaginary quadratic extensions of $\mathbb{Q}$. The latter involves the Galois theory (automorphisms) of the *modular field*, i.e. the field of modular functions associated to the tower of modular curves. This plays an important role in the GL_2 analog of the Bost–Connes system, discussed at length in this book. For all these reasons, it seems appropriate to recall some general facts about Galois theory that illustrate the general philosophy underlying all these different manifestations that are encountered in the book.

We give here a brief account following closely the original exposition by Galois [**136**] and hopefully enticing the reader to think in a fresh way about the general strategy. In particular, we would like to stress the fact that Galois' original formulation shows that Galois theory is in essence describing a symmetry-breaking phenomenon.

Let us start with a ground field k. For simplicity, in this section we only consider the case of characteristic zero, which suffices for our purpose.

We then consider a polynomial equation

$$P(x) = 0 \tag{4.597}$$

of degree n with coefficients in k. We let K denote an algebraic closure of k, so that we can take the roots

$$\mathcal{R} = \{\alpha \in K : P(\alpha) = 0\} \tag{4.598}$$

of (4.597) in K. We assume the roots to be all distinct so that the cardinality of $\mathcal{R}$ is $\#\mathcal{R} = n$. Equivalently, the discriminant of P is $\neq 0$. The discriminant of P is the square of the Vandermonde determinant; it is rational in the coefficients of P.

The first main lemma consists in maximally breaking the symmetry between the roots by choosing an appropriate function V. It was stated and proved by Galois in the following form.

LEMMA 4.210. *Consider a polynomial equation* (4.597) *as above.*

(1) *There exists a rational function* $V(x_1, x_2, \ldots, x_n)$, *of* n *variables, such that* $V(\alpha_1, \ldots, \alpha_n)$ *takes* $n!$ *different values under the permutations of the* n *roots* $\alpha_j \in \mathcal{R} = \{\alpha_1, \ldots, \alpha_n\}$.

(2) *With* V *as above, there exists a rational function* $R(z)$ *with coefficients in* k *such that* $\alpha_1 = R(\xi)$ *for* $\xi = V(\alpha_1, \ldots, \alpha_n)$.

PROOF. (1) We take $V(x_1, x_2, \ldots, x_n) = \sum \lambda_j\, x_j$ with $\lambda_j \in \mathbb{Q} \subset k$. Consider then a non-trivial permutation σ of the points $\{\alpha_1, \ldots, \alpha_n\}$ of $\mathcal{R}$. The equality

$$V(\alpha_1, \ldots, \alpha_n) = V(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)) \tag{4.599}$$

determines a linear subspace $H_\sigma \subset \mathbb{Q}^n$ of codimension at least 1 of the space $\mathbb{Q}^n$ of coefficients λ_j. Since a finite collection of hyperplanes in $\mathbb{R}^n$ cannot cover the dense subset $\mathbb{Q}^n \subset \mathbb{R}^n$, one can choose the λ_j so that none of the equalities (4.599) hold, for the non-trivial permutations $\sigma \in S_n$.
(2) Consider the polynomial

$$Z(X) = \prod_\sigma (X - V(\alpha_1, \sigma(\alpha_2), \ldots, \sigma(\alpha_n))) \tag{4.600}$$

where $\sigma \in S_{n-1}$ ranges through all permutations of the $n-1$ roots $(\alpha_2, \ldots, \alpha_n)$. Since the symmetric functions of these $n-1$ roots are polynomials in α_1 and the symmetric functions s_j of $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, the same holds for all the coefficients $c_k(\alpha_1)$ of the polynomial $Z(X) = \sum c_k\, X^k$. Let $\xi \in K$ be given by $V(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and consider the two equations in x

$$\begin{aligned} &P(x) = 0 \\ &Q(x) = \sum c_k(x)\, \xi^k = 0\,. \end{aligned} \tag{4.601}$$

They are fulfilled by construction for $x = \alpha_1$, but the point is that α_1 is their only common root in the algebraic closure K of k. Indeed, the roots of P are the α_j and one has (with $j = 2$ for instance)

$$\sum c_k(\alpha_2)\, X^k = \prod_\sigma (X - V(\alpha_2, \sigma(\alpha_1), \sigma(\alpha_3), \ldots, \sigma(\alpha_n)))$$

where σ ranges through all permutations of the $n-1$ roots $(\alpha_1, \alpha_3 \ldots, \alpha_n)$. Thus, using the first part of the lemma one gets that $\sum c_k(\alpha_2)\, \xi^k \neq 0$ and α_2 is not a root of the equation $Q(x) = \sum c_k(x)\, \xi^k = 0$.
One then applies the Euclidean algorithm to compute the greatest common divisor of the pair of polynomials P, Q. It is a polynomial of degree 1 whose only root is α_1. This determines α_1 as a rational function of ξ with coefficients that are rational in the s_j and the coefficients λ_j of the linear form V. □

In a similar way one gets $\alpha_2 = R_2(\xi)$, $\alpha_3 = R_3(\xi)$,..., where all these rational functions are different only by the use of permuted values of the coefficients λ_j.
It is interesting to note that the proof of Lemma 4.210 gives in fact a universal form for R as a rational function $R = R(\xi, (s_j), (\lambda_k))$. Its explicit form is very complicated and it is a nice exercise in computer skill to work it out for small values of n. The

fact that explicit computations get quickly out of control is typical of Galois theory, but as Galois insists, there is no point in getting lost in these computations; instead, as he puts it, the goal is

> *"Sauter à pieds joints sur ces calculs; grouper les opérations, les classer suivant leurs difficultés et non suivant leurs formes; telle est selon moi, la mission des géomètres futurs "*
>
> Evariste Galois, [**136**].

He then considers the equation

$$W(x) = \prod_{\sigma}(x - V(\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n))) = 0, \tag{4.602}$$

where $\sigma \in S_n$ ranges through all permutations of the n roots $\{\alpha_1, \ldots, \alpha_n\}$ of (4.597). He then associates to the decomposition

$$W = \prod_i W_i \tag{4.603}$$

of W into irreducible polynomials over k the group G of permutations of the roots of (4.597) given by

$$\begin{array}{cccc} (R_1(\xi), & R_2(\xi), & \cdots, & R_n(\xi)) \\ (R_1(\xi_2), & R_2(\xi_2), & \cdots, & R_n(\xi_2)) \\ & & \cdots & \\ (R_1(\xi_d), & R_2(\xi_d), & \cdots, & R_n(\xi_d)) \end{array} \tag{4.604}$$

where the ξ_j are the roots of the irreducible factor W_1 of (4.603) containing ξ as the root ξ_1.

He explicitly writes ([**136**]) the definition of a group of permutations and characterizes the group of permutations described above in the following form.

THEOREM 4.211. (Galois) *The set of permutations* (4.604) *forms a group G which is uniquely characterized as the only group of permutations of the roots with the following property:*
For any rational fraction $R \in k(x_1, \ldots, x_n)$

$$R((\alpha_j)) \in k \Leftrightarrow R(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)) = R(\alpha_1, \ldots, \alpha_n), \quad \forall \sigma \in G. \tag{4.605}$$

This characterization clearly shows that the group G does not depend upon the explicit choice of the function V in Lemma 4.210.

An equation $E(x) = 0$ such as $E = W_1$ has the peculiar property that the following group of transformations $x \mapsto T(x)$ acts transitively on the roots:

$$\{T \in k(x) \mid E(x) = 0 \Rightarrow E(T(x)) = 0\} \tag{4.606}$$

Indeed, all roots of $W_1(x) = 0$ are of the form $\eta = V(\sigma(\alpha_1), \sigma(\alpha_2), \ldots, \sigma(\alpha_n))$ for some permutation σ while $\alpha_j = R_j(\xi)$ are rational functions of ξ, which allows one to write η as a rational function $T(\xi)$.

We let $F \subset K$ be the field generated by all roots of equation (4.597); then we get:

$$\text{Any root } \eta \text{ of } W(\eta) = 0 \text{ generates } F. \tag{4.607}$$

This shows that, if we decompose W into irreducible factors over k as in (4.603), we then have for each factor W_i an isomorphism

$$F \simeq k[x]/W_i, \quad \forall i. \tag{4.608}$$

In particular all the W_i have the same degree d. Let us then fix W_1 by $W_1(\xi) = 0$ and identify F with $k[\xi]/W_1$. Since all roots of $W = 0$ are rational functions of ξ, the polynomial W_1 has all its roots in F and splits over F as a product of polynomials of degree one. For each root η of $W_1 = 0$ we thus get an automorphism of F

$$\theta \in \mathrm{Aut}_k(F) \tag{4.609}$$

uniquely determined by $\theta(\xi) = \eta$. Their action on the roots of (4.597) is given exactly by (4.604) and Galois' characterization of the group G is simply that $G = \mathrm{Aut}_k(F)$. To be more precise note that if σ is a permutation of the roots such that for any polynomial $Q \in k[x_1, \ldots, x_n]$,

$$Q(\alpha_1, \ldots, \alpha_n) \in k \Rightarrow Q(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)) = Q(\alpha_1, \ldots, \alpha_n)$$

then the map

$$Q(\alpha_1, \ldots, \alpha_n) \mapsto Q(\sigma(\alpha_1), \ldots, \sigma(\alpha_n))$$

gives a well defined homomorphism ρ from F to F (since $Q(\alpha_1, \ldots, \alpha_n) = 0 \in k$ implies $Q(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)) = Q(\alpha_1, \ldots, \alpha_n) = 0$) and ρ is the identity on k so that $\rho \in \mathrm{Aut}_k(F)$. This shows that any group of permutations of the roots fulfilling (4.605) is a subgroup of the group G of permutations of the roots induced by elements of $\mathrm{Aut}_k(F)$. It remains to show that the group $\mathrm{Aut}_k(F)$ fulfills (4.605) and any strict subgroup fails. Given a group H of automorphisms of a field E one lets E^H be the subfield

$$E^H = \{x \in E\,;\ \theta(x) = x \quad \forall \theta \in H\}.$$

The remaining step is immediate using the following well known general result due to E. Artin ([**5**])

Proposition 4.212. *Let E be a field and H a finite group of automorphisms of E.*
(1) *One has $[E : E^H] = \#H$.*
(2) $\mathrm{Aut}_{E^H} E = H$.

Proof. As before, we assume that the fields have characteristic zero. Let us first show that $[E : E^H] \leq \#H$. We let $m = \#H$. We take $m+1$ elements $\xi_0, \ldots, \xi_m \in E$ and consider the matrix

$$T = \begin{pmatrix} \sigma_1(\xi_0) & \ldots & \sigma_1(\xi_m) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \sigma_m(\xi_0) & \ldots & \sigma_m(\xi_m) \end{pmatrix}$$

which has m rows and $m+1$ columns. Thus it has non-trivial kernel and there exists

$$\eta = \begin{pmatrix} \eta_0 \\ \vdots \\ \eta_m \end{pmatrix} \in E^{m+1},$$

such that $T\eta = 0$ and $\eta \neq 0$. One has $\eta_j \neq 0$ for some j and we can assume $j = 0$ and in fact $\eta_0 = 1$ by E-linearity of the kernel. Then $\sum \sigma(\xi_j)\, \eta_j = 0 \Rightarrow \sum \xi_j\, \sigma^{-1}(\eta_j) = 0$. Averaging over H gives a non-trivial relation $\xi_0 + \sum \xi_j\, \lambda_j = 0$, with $\lambda_j \in E^H$. Thus $[E : E^H] \leq \#H$.

Let $k = E^H$. To get a bound on $\mathrm{Aut}_k(E)$ one proves a bound on the cardinality of the set of field morphisms $E \xrightarrow{\theta} K$, where K is an algebraic closure of k. Let $k \subset F \subset F(\xi) \subset E$ be subfields, with $\xi \notin F$. Then let P be the monic irreducible polynomial with coefficients in F such that $P(\xi) = 0$ (the minimal polynomial of ξ). Its degree is $[F(\xi) : F]$.
If we know θ on F, then to extend it to $F(\xi)$ requires choosing a solution of $\theta(P)(X) = 0$ in K and there are $\deg P = [F(\xi) : F]$ such solutions. Thus, writing $k \subset k(\xi_1) \subset k(\xi_1)(\xi_2) \subset \cdots = E$, one gets the bound

$$\mathrm{Card}(\mathrm{Hom}_k(E, K)) \leq [E : k]. \tag{4.610}$$

It follows thus that $\#\mathrm{Aut}_k(E) \leq [E : k]$ and (1) and (2) follow. $\square$

Corollary 4.213. *Let E and H be as above, with $k = E^H$. Then, for $\xi \in E$,*

$$\xi \text{ generates } E \text{ over } k \Leftrightarrow \theta(\xi) \neq \xi \quad \forall\, \theta \neq 1 \quad (\theta \in H)\,.$$

Proof. The implication $\Rightarrow$ is immediate. To get the other, let P be the minimal polynomial of ξ. Then all $\theta(\xi)$ are roots of P and are distinct, so that $\deg(P) \geq m = \#H$. The subfield $k(\xi)$ has degree m over k and equals E. $\square$

Proposition 4.214. *Let E, H, and $k = E^H$ be as above. Let $\xi_j \in E$ be such that if $\theta(\xi_j) = \xi_j$, for all j, then $\theta = 1$. It follows that E is the ring generated by the ξ_j over k.*

Proof. Let R be the ring $k[\xi_j]$. Since we have $R \subset E$, then every non-zero element $\eta \in R$ is injective as a multiplier of R, hence also surjective on this finite dimensional k-vector space. Thus R is a subfield of E and it is enough to show that the subfield generated by the ξ_j is E.
Let $k_1 = k(\xi_1)$ and P_1 be the corresponding irreducible polynomial of degree n_1. Let

$$H_1 = \{\theta \in H \;:\; \theta(\xi_1) = \xi_1\}$$

and $m_1 = \#H_1$, with $m_1 \mid m$. Then consider

$$P_1(X) = \prod_{H_1 \backslash H} (X - \theta(\xi_1)) \tag{4.611}$$

as above. Thus $\deg k_1 = \deg P_1 = m/m_1$. Since $[E : E^{H_1}] = m_1$ and $k_1 \subset E^{H_1}$, we get

$$k_1 = E^{H_1}\,. \tag{4.612}$$

We can then go on with ξ_2 and $k_2 = k_1(\xi_2)$, so that $k_2 = E^{H_2}$ with

$$H_2 = \{\theta \in H_1 \;:\; \theta(\xi_2) = \xi_2\}\,. \tag{4.613}$$

In the end we get $H_n = \{1\}$ with $k_n = E = k(\xi_1, \ldots, \xi_n)$ as required. $\square$

In fact the proof above applies without change and yields the "fundamental theorem" of Galois theory.

Theorem 4.215. *With E, H, and $k = E^H$ as above, consider elements $\xi_j \in E$. Then the subring generated by the ξ_j is the fixed subfield E^J with*

$$J = \{\theta \in H\,;\; \theta(\xi_j) = \xi_j \quad \forall\, j\}\,.$$

Thus, given any subfield $k \subset F \subset E$, one has

$$F = E^J \,, \quad \text{with} \quad J = \{\theta \in H \,:\, \theta(\xi) = \xi \quad \forall \xi \in F\}\,, \tag{4.614}$$

and one gets a bijection between subgroups of H and subfields of E. Normal subgroups correspond to *normal extensions*, *i.e.* one has the following result.

PROPOSITION 4.216. *Let $k \subset F$ be a finite extension. Then the following are equivalent.*

(1) *$k = F^G$ where $G = \mathrm{Aut}_k(F)$.*
(2) *Any irreducible polynomial $P \in k[X]$ with one root in F has all its roots in F.*
(3) *There exists $\xi \in F$, with $k(\xi) = F$, such that all roots of $P(\xi) = 0$ (minimal polynomial of ξ) are in F.*

PROOF. (1) $\Rightarrow$ (2). Let α be that root. Then $F_1 = k(\alpha)$ is a subfield of degree $n_1 = \deg P$ and is of the form F^{H_1}, for $H_1 = \{\theta \in G \,:\, \theta(\alpha) = \alpha\}$. This implies that the $\theta(\alpha)$, for $\theta \in H_1 \backslash G$, are the n_1 roots of P in F and P splits in F.
(2) $\Rightarrow$ (3). Take elements $\alpha_i \in F$ that generate F over k and corresponding irreducible polynomials $P_i \in k[X]$. Then $P = \prod_i P_i$ has all its roots in F by (2). By Lemma 4.210 (which does not use irreducibility), we find a rational function ξ of the roots which generates F over k. Taking the minimal polynomial of ξ and using (2) we get (3).
(3) $\Rightarrow$ (1). One constructs $\deg(P)$ automorphisms of F over k. $\square$

If $J \lhd H$ is a *normal* subgroup, then the quotient group H/J acts on $E^J = F$ as automorphisms, with k as fixed points. Thus, F fulfills (1) of Proposition 4.216 above. Conversely, if F fulfills (3) of Proposition 4.216, then the subgroup $H_1 = \{\theta \in G \,:\, \theta(\xi) = \xi\}$ is unchanged by replacing ξ with $\alpha(\xi)$, for $\alpha \in G$, since $\alpha(\xi) \in F$ is another root of P. Thus H_1 is a normal subgroup of G. In other words, in the Galois correspondence between subfields and subgroups, normal subgroups correspond to normal extensions.

This brief excursion into the basic facts of Galois theory serves as a quick introduction to the many different variants of the theory that are discussed in the book. As a fundamental tool in number theory, it appears repeatedly in our encounters between noncommutative geometry and arithmetic.

Bibliography

[1] S.L. Adler, *Axial-vector vertex in spinor electrodynamics*, Phys. Rev., Vol.177 (1969) 2426–2438.

[2] Y. André, *Une introduction aux motifs*, Panoramas et Synthèses, Vol.17, Société mathématique de France, 2005.

[3] H. Araki, *Expansional in Banach algebras*, Ann. Sci. École Norm. Sup. (4) 6 (1973), 67–84.

[4] H.Arason, D.J. Castano, B. Kesthlyi, E.J.Piard, P.Ramond, B.D.Wright, *Renormalization-group study of the standard model and its extensions: the standard model*, Phys. Rev. D, 46 (1992) N.9, 3945–3965.

[5] E. Artin, *Galois theory.* Edited and with a supplemental chapter by Arthur N. Milgram. Reprint of the 1944 second edition. Dover Publications, Inc., Mineola, NY, 1998.

[6] E. Artin, J. Tate, *Class field theory.* W. A. Benjamin, 1968.

[7] A. Ashtekar, *Lectures on non-perturbative canonical gravity*, World Scientific, 1991.

[8] M.F. Atiyah. *K-Theory.* Benjamin, New York, 1967.

[9] M.F. Atiyah, R. Bott, *A Lefschetz fixed point formula for elliptic complexes: I*, Annals of Math. Vol. 86 (1967) 374–407.

[10] M.F. Atiyah, F. Hirzebruch, *Vector bundles and homogeneous spaces.* 1961 Proc. Sympos. Pure Math., Vol. III pp. 7–38 American Mathematical Society.

[11] I. G. Avramidi *Covariant methods for the calculation of the effective action in quantum field theory and investigation of higher-derivative quantum gravity.* PhD Thesis, Moscow University 1986, hep-th/9510140.

[12] S. Baaj, P. Julg, *Théorie bivariante de Kasparov et opérateurs non bornés dans les C^*-modules hilbertiens.* C. R. Acad. Sci. Paris Sér. I Math. 296 (1983), no. 21, 875–878.

[13] J. W. Barrett *A Lorentzian version of the non-commutative geometry of the standard model of particle physics*, hep-th/0608221

[14] P Belkale, P. Brosnan, *Matroids, motives, and a conjecture of Kontsevich.* Duke Math. J. 116 (2003), no. 1, 147–188.

[15] J.S. Bell, R. Jackiw, *A PCAC puzzle*, Nuovo Cimento, Vol.60A (1969) 47–60.

[16] M-T. Benameur, T. Fack, *Type II non-commutative geometry. I. Dixmier trace in von Neumann algebras.* Adv. Math. 199 (2006), no. 1, 29–87.

[17] J. Bernstein, *The analytic continuation of generalized functions with respect to a parameter*, Funct. Anal. Appl. Vol. 6 no 4 (1972) 273–285.

[18] M. Berry, *Riemann's zeta function: a model of quantum chaos*, Lecture Notes in Physics, Vol.263, Springer-Verlag, 1986.

[19] D. Bessis, C. Itzykson, J.B. Zuber, *Quantum field theory techniques in graphical enumeration*, Adv. Appl. Math. 1 (1980) 109–157.

[20] B. Blackadar, *K-theory for operator algebras*, Second edition, MRSI Publications, 1998.

[21] S. Bloch, *Lectures of mixed motives*, in "Algebraic geometry—Santa Cruz 1995", pp.329–359, Proc. Sympos. Pure Math., 62, Part 1, Amer. Math. Soc., 1997.

[22] S. Bloch, *Algebraic cycles and higher K-theory*, Adv. in Math. 61 (1986), no. 3, 267–304.

[23] S. Bloch, *Function theory of polylogarithms. Structural properties of polylogarithms*, 275–285, Math. Surveys Monogr., 37, Amer. Math. Soc., Providence, RI, 1991.

[24] S. Bloch, *Motives associated to graphs*, Japanese Journal of Mathematics, (2007) N.2, 165–196.

[25] S. Bloch, H. Esnault, D. Kreimer, *On motives associated to graph polynomials*, Comm. Math. Phys. 267 (2006), N.1, 181–225.

[26] N.N. Bogoliubov, O. Parasiuk, *Über die Multiplikation der Kausal Funktionen in der Quantentheorie der Felder*, Acta Mathematica 97 (1957) 227–266.

[27] N.N. Bogoliubov and D.V. Shirkov, *Introduction to the theory of quantized fields*, 3rd ed., Wiley 1980.

[28] C. Bordé, *Base units of the SI, fundamental constants and modern quantum physics*, Phil. Trans. R. Soc. A 363 (2005) 2177–2201.

[29] J.-B. Bost and A. Connes, *Produits Euleriens et facteurs de type III*, C. R. Acad. Sci. Paris Sér. I Math. Vol.315 (1992), no. 3, 279–284.

[30] J.B. Bost, A. Connes, *Hecke algebras, Type III factors and phase transitions with spontaneous symmetry breaking in number theory*, Selecta Math. (New Series) Vol.1 (1995) N.3, 411–457.

[31] G.D Boyd, J.P. Gordon, *Confocal multimode resonator for millimeter through optical wavelength masers*, Bell Syst. Tech. Journal, Vol.40 (1961), 489–508.

[32] S. Bradlow, *Vortices in holomorphic line bundles over closed Kähler manifolds*, Comm. Math. Phys., Volume 135, Number 1 (1990) 1–17.

[33] O. Bratteli, D.W. Robinson, *Operator algebras and quantum statistical mechanics I,II*, Springer Verlag, 1979 and 1981.

[34] P. Breitenlohner, D. Maison, *Dimensional renormalization and the action principle*. Comm. Math. Phys. Vol.52 (1977), N.1, 11–38.

[35] D.J. Broadhurst, D. Kreimer, *Association of multiple zeta values with positive knots via Feynman diagrams up to* 9 *loops*. Phys. Lett. B 393 (1997), no. 3-4, 403–412.

[36] F. Bruhat, *Distributions sur un groupe localement compact et applications à l'étude des représentations des groupes p-adiques*. Bull. Soc. Math. France, 89 (1961), 43–75.

[37] J.F. Burnol, *The explicit formula and the conductor operator*, math.NT/9902080.

[38] K.Cahill, *The strong CP problem in a compact Robertson-Walker universe*, Z. Phys. C Vol.29 (1985) 537-539.

[39] Alan L. Carey, J. Phillips, A. Rennie, F. A. Sukochev *The local index formula in semifinite von Neumann algebras. I. Spectral flow*, Adv. Math. 202 (2006), no. 2, 451–516.

[40] Alan L. Carey, J. Phillips, A. Rennie, F. A. Sukochev *The local index formula in semifinite von Neumann algebras. II. The even case*, Adv. Math. 202 (2006), no. 2, 517–554.

[41] L. Carminati, B. Iochum, T. Schucker, *Noncommutative Yang–Mills and noncommutative relativity: a bridge over trouble water*, Eur.Phys.J. C8 (1999) 697–709.

[42] P. Cartier, *Grothendieck et les motifs*, preprint IHES/M/00/75.

[43] J.A. Casas, J.R. Espinosa, A. Ibarra, I. Navarro, *General RG equations for physical neutrino parameters and their phenomenological implications*, Nucl.Phys. B573 (2000) 652–684.

[44] CDF collaboration, *Observation of* B_s^0–$\bar{B}_s^0$ *oscillations*, arXiv hep-ex/0609040.

[45] A. Chamseddine, A. Connes, *Universal formula for noncommutative geometry actions: unification of gravity and the Standard Model*, Phys. Rev. Lett. 77, 486804871 (1996).

[46] A. Chamseddine, A. Connes, *The Spectral action principle*, Comm. Math. Phys. Vol.186 (1997), 731–750.

[47] A. Chamseddine, A. Connes, *Scale invariance in the spectral action*, J. Math. Phys. 47 (2006), N.6, 063504, 19 pp.

[48] A. Chamseddine, A. Connes, *Inner fluctuations of the spectral action*, J. Geom. Phys. 57 (2006), N.1, 1–21.

[49] A. Chamseddine, A. Connes, *Why the Standard Model*, hep-th 0706.3688.

[50] A. Chamseddine, A. Connes, *A dress for SM the beggar*, hep-th 0706.3690.

[51] A. Chamseddine, A. Connes, *Quantum gravity boundary terms from the spectral action of noncommutative space*, Phys. Rev. Letter LE11230, hep-th 0705.1786.

[52] A. Chamseddine, A. Connes, M. Marcolli, *Gravity and the standard model with neutrino mixing*, hep-th/0610241.

[53] D. Chang, R.N. Mohapatra, J.M. Gipson, R.E. Marshak, M.K. Parida, *Experimental tests of new* $SO(10)$ *grand unification*, Phys. Rev. D 31 (1985) N.7, 1718-1732.

[54] V. Chari, A.N. Pressley, *A Guide to Quantum Groups*, Cambridge University Press, 1994.

[55] K.T. Chen, *Iterated integrals and exponential homomorphisms*, Proc. London Math. Soc. Vol.3 N.4 (1954), 502–512.

[56] K.T. Chen, *Iterated integrals of differential forms and loop space homology*. Ann. of Math. (2) 97 (1973), 217–246.

[57] T.P. Cheng and L.F. Li, *Gauge theory of elementary particle physics*, Oxford, 1984.

[58] A. Codello, R. Percacci, *Fixed points of higher derivative gravity*, Phys.Rev.Lett. 97 (2006) 221301.

[59] S. Coleman, *Lectures on Quantum Field Theory*, unpublished manuscript.

[60] S. Coleman, *Renormalization and symmetry: a review for non-specialists*, in “Aspects of Symmetry” pp.99–112, Cambridge University Press, 1985.

[61] S. Coleman, *Aspects of symmetry*, Selected Erice Lectures, Cambridge University Press, 1985.

[62] J. Collins, *Renormalization*, Cambridge Monographs in Math. Physics, Cambridge University Press, 1984.

[63] A. Connes, *Une classification des facteurs de type* III. Ann. Sci. École Norm. Sup. (4) 6 (1973), 133–252.

[64] A. Connes, *C^* algèbres et géométrie différentielle*, C. R. Acad. Sci. Paris Sér. A-B 290 (1980), no. 13, A599–A604.

[65] A. Connes, *Cohomologie cyclique et foncteurs* Ext^n. C. R. Acad. Sci. Paris Sér. I Math. 296 (1983), no. 23, 953–958.

[66] A. Connes, *A survey of foliations and operator algebras.* In “Operator algebras and applications”, Part I (Kingston, Ont., 1980), pp. 521–628, Proc. Sympos. Pure Math., 38, Amer. Math. Soc., Providence, R.I., 1982.

[67] A. Connes, *Noncommutative differential geometry.* Inst. Hautes Études Sci. Publ. Math. No. 62 (1985), 257–360.

[68] A. Connes, *Noncommutative geometry*, Academic Press (1994).

[69] A. Connes, *Non commutative geometry and reality*, Journal of Math. Physics 36 no. 11 (1995).

[70] A. Connes, *Gravity coupled with matter and the foundation of noncommutative geometry*, Comm. Math. Phys. (1995)

[71] A. Connes, *Trace formula in noncommutative geometry and the zeros of the Riemann zeta function.* Selecta Math. (N.S.) 5 (1999), no. 1, 29–106.

[72] A. Connes, *Formules explicites, formules de trace et réalisation spectrale des zéros de la fonction zéta*, Course at Collège de France, 1999.

[73] A. Connes, *Noncommutative Geometry and the standard model with neutrino mixing*, J. High Energy Phys. 2006, no. 11, 081, 19 pp.

[74] A. Connes, C. Consani, M. Marcolli, *Noncommutative geometry and motives: the thermodynamics of endomotives*, Advances in Mathematics, Vol.214 (2007) N.2, 761–831.

[75] A. Connes, C. Consani, M. Marcolli, *Weil's proof and the geometry of the adeles class space*, math/0703392.

[76] A. Connes, J. Cuntz, *Quasi homomorphismes, cohomologie cyclique et positivité*, Comm. Math. Phys. 114 (1988).

[77] A. Connes, M. Douglas, A. Schwarz, *Noncommutative geometry and matrix theory: compactification on tori.* J. High Energy Phys. 1998, no. 2, Paper 3, 35 pp. (electronic).

[78] A. Connes, M. Dubois-Violette, *Noncommutative finite-dimensional manifolds. I. Spherical manifolds and related examples.* Comm. Math. Phys. 230 (2002), no. 3, 539–579.

[79] A. Connes, M. Dubois-Violette, *Moduli space and structure of noncommutative 3-spheres.* Lett. Math. Phys. 66 (2003), no. 1-2, 91–121.

[80] A. Connes, M. Dubois-Violette, *Non commutative finite-dimensional manifolds II. Moduli space and structure of noncommutative 3-spheres.* ArXiv math.QA/0511337.

[81] A. Connes, M. Karoubi, *Caractère multiplicatif d'un module de Fredholm.* K-Theory 2 (1988), no. 3, 431–463.

[82] A. Connes, D. Kreimer, *Renormalization in quantum field theory and the Riemann-Hilbert problem. I. The Hopf algebra structure of graphs and the main theorem.* Comm. Math. Phys. 210 (2000), no. 1, 249–273.

[83] A. Connes, D. Kreimer, *Renormalization in quantum field theory and the Riemann-Hilbert problem. II. The β-function, diffeomorphisms and the renormalization group.* Comm. Math. Phys. 216 (2001), no. 1, 215–241.

[84] A. Connes and G. Landi, *Noncommutative manifolds, the instanton algebra and isospectral deformations.* Comm. Math. Phys. 221 (2001), no. 1, 141–159.

[85] A. Connes, J. Lott, *Particle models and noncommutative geometry.* Recent advances in field theory (Annecy-le-Vieux, 1990). Nuclear Phys. B Proc. Suppl. 18B (1990), 29–47 (1991).

[86] A. Connes, M. Marcolli, *From physics to number theory via noncommutative geometry, Part I: Quantum statistical mechanics of* $\mathbb{Q}$*-lattices*, in "Frontiers in Number Theory, and Geometry, I" pp.269–350, Springer Verlag, 2006.

[87] A. Connes, M. Marcolli, *From Physics to Number theory via Noncommutative Geometry, II: Renormalization, the Riemann–Hilbert correspondence, and motivic Galois theory*, to appear in "Frontiers in Number Theory, Physics, and Geometry, II" pp.617–713, Springer Verlag, 2006.

[88] A. Connes, M. Marcolli, $\mathbb{Q}$*-lattices: quantum statistical mechanics and Galois theory*, J. Geom. Phys. 56 (2005) N.1, 2–25.

[89] A. Connes, M. Marcolli, *Renormalization and motivic Galois theory*, International Math. Research Notices, (2004), no. 76, 4073–4091.

[90] A. Connes, M. Marcolli, N. Ramachandran, *KMS states and complex multiplication*, Selecta Math. (New Ser.) Vol.11 (2005), no. 3-4, 325–347.

[91] A. Connes, M. Marcolli, N. Ramachandran *KMS states and complex multiplication. Part II*, in "Operator Algebras: The Abel Symposium 2004", pp.15–59, Abel Symp., 1, Springer Verlag, 2006.

[92] A. Connes, H. Moscovici, *The local index formula in noncommutative geometry*, GAFA, Vol. 5 (1995), 174–243.

[93] A. Connes, H. Moscovici, *Modular Hecke Algebras and their Hopf Symmetry*, Mosc. Math. J. 4 (2004), no. 1, 67–109, 310.

[94] A. Connes, H. Moscovici, *Rankin-Cohen Brackets and the Hopf Algebra of Transverse Geometry*, Mosc. Math. J. 4 (2004), no. 1, 111–130, 311.

[95] A. Connes, H. Moscovici, *Transgressions of the Godbillon-Vey class and Rademacher functions*, in "Noncommutative Geometry and Number Theory", pp.79–107. Vieweg Verlag, 2006.

[96] A. Connes, M. Rieffel, *Yang-Mills for noncommutative two-tori.* Operator algebras and mathematical physics (Iowa City, Iowa, 1985), 237–266, Contemp. Math., 62, Amer. Math. Soc., Providence, RI, 1987.

[97] A. Connes and C. Rovelli. *von Neumann algebra automorphisms and time-thermodynamics relation in generally covariant quantum theories.* Classical Quantum Gravity, 11 (1994) N.12, 2899–2917.

[98] A. Connes, G. Skandalis, *The longitudinal index theorem for foliations*, Publ. RIMS Kyoto Univ. 20 (1984) 1139–1183.

[99] A. Connes, M. Takesaki, *The flow of weights on factors of type III.* Tohoku Math. J., 29, (1977) 473–575.

[100] C. Consani, *Double complexes and Euler L-factors.* Compositio Math. 111 (1998), no. 3, 323–358.

[101] C. Consani, C. Faber, *On the cusp form motives in genus 1 and level 1*, math.AG/0504418.

[102] C. Consani, M. Marcolli, *Noncommutative geometry, dynamics, and* ∞*-adic Arakelov geometry.* Selecta Math. (N.S.) 10 (2004), no. 2, 167–251.

[103] C. Consani, M. Marcolli, *New perspectives in Arakelov geometry.* Number theory, 81–102, CRM Proc. Lecture Notes, 36, Amer. Math. Soc., Providence, RI, 2004.

[104] C. Consani, M. Marcolli, *Quantum statistical mechanics over function fields*, math.QA/0607363.

[105] P. Deligne, *Equations differentielles à points singuliers réguliers*, Lecture Notes in Mathematics 163, Springer 1970.

[106] P. Deligne, *Catégories tannakiennes*, in "Grothendieck Festschrift" Vol.2, pp. 111–195, Progress in Mathematics Vol.87, Birkhäuser, 1990.

[107] P. Deligne, *Structures de Hodge mixtes relles*, Motives (Seattle, WA, 1991), 509–514, Proc. Sympos. Pure Math., 55, Part 1, Amer. Math. Soc., Providence, RI, 1994.

[108] P. Deligne, *A quoi servent les motifs?*, in "Motives" (Seattle, WA, 1991), pp.143–161, Proc. Sympos. Pure Math., 55, Part 1, Amer. Math. Soc., 1994.

[109] P. Deligne, *Valeurs de fonctions L et périodes d'intégrales*, Proc. Symp. Pure Math. Vol.33 (1979) part II, 313–346.

[110] P. Deligne, *La conjecture de Weil. I.* Publ. Math. Inst. Hautes Études Sci. No. 43 (1974), 273–307.

[111] P. Deligne, *Théorie de Hodge III*, Pub. Math. IHES 44 (1975) 6–77.

[112] P. Deligne, *Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques.* in "Automorphic forms, representations and L-functions" Part 2, pp. 247–289, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., 1979.

[113] P. Deligne, A.B. Goncharov, *Groupes fondamentaux motiviques de Tate mixte*, math.NT/0302267.

[114] P. Deligne, J.S. Milne, , in "Hodge cycles, motives, and Shimura varieties", Lecture Notes in Mathematics, 900, pp.101–228. Springer-Verlag, 1982

[115] M. Demazure, A. Grothendieck, et al. *Séminaire Géometrie Algébrique: Schémas en Groupes*, Lecture Notes in Mathematics, Vol. 151, 152, 153. Springer, 1970.

[116] C. Deninger, *On the Γ-factors attached to motives*, Invent. Math. 104 (1991) 245–261.

[117] C. Deninger, *Motivic L-functions and regularized determinants*, in "Motives", Proceedings of Symposia in Pure Mathematics, Vol. 55 (1994) Part I, 707–743.

[118] C. Deninger, *On the nature of the "explicit formulas" in analytic number theory—a simple example.* in "Number theoretic methods" (Iizuka, 2001), 97–118, Dev. Math., 8, Kluwer Acad. Publ., 2002.

[119] P. Dirac, *The quantum theory of the emission and absorption of radiation.* Proc. London Royal Soc. 114 (1927) 243–265.

[120] J. Dixmier, *Les C*-algèbres et leurs representations* Reprint of the second (1969) edition. Les Grands Classiques Gauthier-Villars. Éditions Jacques Gabay, Paris, 1996. 403 pp.

[121] S.K. Donaldson, P.B. Kronheimer, *The geometry of four-manifolds*, Oxford Univ. Press, 1990.

[122] J.F. Donoghue, *General relativity as an effective field theory: the leading quantum corrections*, Phys. Rev. D, Vol.50 (1994) N.6, 3874–3888.

[123] J.F. Donoghue, E. Golowich, B.R. Holstein, *Dynamics of the Standard Model*, Cambridge University Press, 1994.

[124] S. Doplicher, J.E. Roberts, *A new duality for compact groups.* Invent. Math. 98 (1989) 157–218.

[125] M. Dresden, *Renormalization in historical perspective - The first stage*, in "Renormalization", ed. L. Brown, Springer-Verlag, 1994.

[126] H.M. Edward, *Riemann's zeta function*, Dover, 2001.

[127] M.B. Einhorn, D.R.T. Jones, *Effective potential and quadratic divergences*, Phys. Rev. D, Vol.46 (1992) N.11, 5206-5208.

[128] P. Etingof, *Note on dimensional regularization*, in "Quantum Fields and Strings: a Course for Mathematicians", Vol.I, pp. 595–607. American Mathematical Society, 1999.

[129] P. Falloon, *Theory and Computation of Spheroidal Harmonics with General Arguments*, Ph.D. Thesis, University of Western Australia, 2001.

[130] R. Feynman, *The reason for antiparticles*, in "Elementary Particles and the Laws of Physics", Cambridge Univ. Press (1987).

[131] R. Feynman, *Feynman lectures on gravitation*, Perseus books (1995).

[132] H. Figueroa, J.M. Gracia-Bondía, J. Varilly, *Elements of Noncommutative Geometry*, Birkhäuser, 2000.

[133] A.G. Fox, T. Li, *Resonant modes in a maser interferometer*, Bell Syst. Tech. Journal, Vol.40 (1961), 453–458.

[134] P.H. Frampton, *Conditions for renormalizability of quantum flavor dynamics*, Phys. Rev. D, Vol.20 (1979) 33723377

[135] J. Fröhlich, K. Gawedski, *Conformal field theory and geometry of strings*, in "Mathematical quantum theory. I. Field theory and many-body theory" pp. 57–97, CRM Proc. Lecture Notes, 7, AMS, 1994.

[136] E. Galois, Oeuvres de Galois, Gauthier-Villars, Paris (1962).

[137] O. García-Prada, *A direct existence proof for the vortex equations over a compact Riemann surface.* Bull. London Math. Soc. 26 (1994), no. 1, 88–96.

[138] I. M. Gelfand, N. J. Vilenkin, *Generalized Functions Vol IV*, Academic press, New-York, London, San-Diego(1964)

[139] S.I. Gelfand, Yu.I. Manin, *Homological algebra*, Encyclopaedia of Mathematical Sciences, Vol.38, Springer, 1994.

[140] C.Q. Geng, R.E. Marshak, *Uniqueness of quark and lepton representations in the standard model from the anomalies viewpoint*, VPI-IHEP-88/5, 14pp.

[141] H. Georgi and S.L. Glashow, *Gauge Theories without anomalies*, Phys. Rev. D, Vol.6 (1972) 429–431.

[142] E. Getzler, M. Kapranov, *Modular operads*. Compositio Math. 110 (1998), no. 1, 65–126.

[143] S. Giddings, D. Marolf, J. Hartle, *Observables in effective gravity*, Phys. Rev. D 74, 064018 (2006) 20 pp.

[144] P. Gilkey, *Invariance Theory, the Heat Equation and the Atiyah-Singer Index Theorem*, Wilmington, Publish or Perish, 1984.

[145] J. Glimm and A. Jaffe. *Quantum physics*. Springer, 1981

[146] A. Goncharov, *Multiple polylogarithms and mixed Tate motives*, math.AG/0103059.

[147] A. Goncharov, *Volumes of hyperbolic manifolds and mixed Tate motives*. J. Amer. Math. Soc. 12 (1999), no. 2, 569–618.

[148] A. Goncharov, Yu.I. Manin, *Multiple ζ-motives and moduli spaces $\overline{\mathcal{M}}_{0,n}$*. Compos. Math. 140 (2004), no. 1, 1–14.

[149] J. Gracia-Bondía, B. Iochum, T. Schucker, *The standard model in noncommutative geometry and fermion doubling*. Phys. Lett. B 416 no. 1-2 (1998), 123–128.

[150] G. Green, *Researches on the Vibrations of Pendulums in Fluid Media*, Royal Society of Edinburgh Transactions (1836) 315–324.

[151] D. Gross, *Applications of the renormalization group to high energy physics*, in "Methods In Field Theory", Les Houches 1975, Proceedings, Amsterdam, (1976), 141–250.

[152] D. Gross and R. Jackiw, *Effect of anomalies on quasi-renormalizable theories*, Phys. Rev. D, Vol.6 (1972) 477–493.

[153] A. Grothendieck, *Sur la classification des fibrés holomorphes sur la sphère de Riemann*, Amer. J. Math. Vol.79 (1957) 121–138.

[154] A. Grothendieck, et al. *Revêtements étales et groupe fondamental*, Lecture Notes in Mathematics, 224, Springer, 1971. (Available electronically as math.AG/0206203)

[155] A. Grothendieck, *Standard conjectures on algebraic cycles*, in "Bombay coll. on Alg. Geom." Oxford University Press, 1969, pp.193–199.

[156] V. Guillemin, S. Sternberg, *Geometric asymptotics*, Math. Surveys Vol. 14, American Mathematical Society, 1977.

[157] V. Guillemin, *Lectures on spectral theory of elliptic operators*, Duke Math. J., Vol. 44, 3 (1977), 485-517.

[158] L. Guo, B. Zhang, *Renormalization of multiple zeta values*, preprint math.NT/0606076.

[159] E. Ha, F. Paugam, *Bost-Connes-Marcolli systems for Shimura varieties. I. Definitions and formal analytic properties*, IMRP Int. Math. Res. Pap. 2005, no. 5, 237–286.

[160] R. Haag, *Local Quantum Physics*, Springer, 1992.

[161] R. Haag, N. M. Hugenholtz, M. Winnink *On the equilibrium states in quantum statistical mechanics*, Comm. Math. Phys. 5 (1967), 215, 236.

[162] M. Hanamura, *Mixed motives and algebraic cycles. I.* Math. Res. Lett. 2 (1995), no. 6, 811–821.

[163] S. Haran, *Riesz potentials and explicit sums in arithmetic*, Invent. Math., 101 (1990), 697-703.

[164] J. Hartle, S. Hawking, *Wave function of the universe*. Phys. Rev. D 28, 2960 (1983)

[165] S. Hawking, *Hawking on the big bang and black holes*. Advanced Series in Astrophysics and Cosmology, Vol. 8. World Scientific (1993).

[166] K. Hepp, *Proof of the Bogoliubov-Parasiuk theorem on renormalization*, Comm. Math. Phys. 2, (1966), 301–326.

[167] W.F.L. Hollik, *Radiative corrections in the Standard Model*, Fortschritte der Physik, Vol.38 (1990) N.3 165–260.

[168] G. 't Hooft, *Renormalization of massless Yang-Mills fields*, Nucl. Phys. B, 35 (1971) 167–188.

[169] G. 't Hooft, M. Veltman, *Regularization and renormalization of gauge fields* Nuclear Physics B, Vol.44, N.1 (1972), 189–213.

[170] G. 't Hooft, M. Veltman, *Diagrammar*, CERN Report 73/9 (1973), reprinted in "Particle interactions at very high energies". NATO Adv. Study Inst. Series, Sect. B, vol. 4B, 177.

[171] T. Inagaki, H. Suzuki, *Majorana and Majorana–Weyl fermions in lattice gauge theory*, JHEP 07 (2004) 038, 30 pages.

[172] B. Jacob, *Bost-Connes type systems for function fields*, J. Noncommutative Geometry, Vol.1 (2007) N.2, 141–211.

[173] U. Jannsen, S. Kleiman and J.P. Serre (Eds.) *Motives.* Proceedings of the AMS-IMS-SIAM Joint Summer Research Conference held at the University of Washington, Seattle, Washington, July 20–August 2, 1991. Proceedings of Symposia in Pure Mathematics, 55, Part 1. American Mathematical Society, Providence, RI, 1994.

[174] U. Jannsen, *Motives, numerical equivalence, and semi-simplicity.* Invent. Math. 107 (1992), no. 3, 447–452.

[175] G. Janusz, *Algebraic number fields*, Academic Press, 1973.

[176] F. Jegerlehner, *Facts of life with gamma(5),* Eur. Phys. J. C, Vol.18 (2001) N.4, 673–679.

[177] G. Kane, *Modern elementary particle physics*, Addison-Wesley, 1987.

[178] G.G. Kasparov, *The operator K-functor and extensions of C^*-algebras.* Izv. Akad. Nauk SSSR Ser. Mat. 44 (1980), no. 3, 571–636, 719.

[179] D. Kastler, *The Dirac operator and gravitation*, Commun. Math. Phys. 166 (1995), 633–643.

[180] D. Kastler, *Noncommutative geometry and fundamental physical interactions: The Lagrangian level*, Journal. Math. Phys. 41 (2000), 3867–3891.

[181] D. Kastler, *A detailed account of Alain Connes' version of the standard model in noncommutative geometry. I, II.* Rev. Math. Phys. 5 (1993), no. 3, 477–532.

[182] D. Kastler, *Introduction to algebraic theory of superselection sectors (space-time dimension $= 2$, strictly localized morphisms).* in "The algebraic theory of superselection sectors" (Palermo, 1989), pp. 113–214, World Scientific, 1990.

[183] T. Kato, *Perturbation theory of linear operators* Grundlerhren der Mathematischen, Vol.132, Springer, 1966.

[184] R. Kirchner, I. Montvay, J. Westphalen, S. Luckmann, K. Spanderen, *Evidence for discrete chiral symmetry breaking in N=1 supersymmetric Yang-Mills theory*, hep-lat/9810062

[185] S.L. Kleiman, *Algebraic cycles and the Weil conjectures*, in "Dix exposés sur la cohomologie des schémas", North Holland, 1968, pp.359–386.

[186] S.L. Kleiman, *Motives*, in "Algebraic geometry, Oslo 1970" (Proc. Fifth Nordic Summer-School in Math., Oslo, 1970), pp. 53–82. Wolters-Noordhoff, 1972.

[187] M. Knecht, T. Schucker *Spectral action and big desert* Phys. Lett. B 640 (2006) 272–277.

[188] I.Yu. Kobzarev, Yu.I. Manin, *Elementary Particles: Mathematics, Physics and Philosophy*, Kluwer, 1989.

[189] M. Kontsevich, Yu.I. Manin, *Gromov–Witten classes, quantum cohomology and enumerative geometry*, Comm. Math. Phys. 164 (1994) 525–562.

[190] D. Kreimer, *Residues of quantum fields – numbers we should know*, in "Noncommutative Geometry and Number Theory", pp.187–203, Aspects of Mathematics Vol.37, Vieweg Verlag, 2006.

[191] M. Laca, *Semigroups of $*$-endomorphisms, Dirichlet series, and phase transitions*, J. Funct. Anal. 152 (1998) 330–378.

[192] M. Laca, *From endomorphisms to automorphisms and back: dilations and full corners*, J. London Math. Soc. (2) 61 (2000), no. 3, 893–904.

[193] M. Laca, N. Larsen, S. Neshveyev, *Phase transitions in the Connes–Marcolli* GL_2*-system*, preprint 2006.

[194] M. Laca, M. van Frankenhuijsen, *Phase transitions on Hecke C^*-algebras and class-field theory over* $\mathbb{Q}$, J reine angew. Math. 595 (2006) 25–53.

[195] M. Laca and I. Raeburn, *Semigroup crossed products and the Toeplitz algebras of nonabelian groups*, J. Funct. Anal., **139** (1996), 415–440.

[196] H.J. Landau, H. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty II*, Bell Syst. Tech. J. Vol.40 (1961).

[197] H.J. Landau, H. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty III*, Bell Syst. Tech. J. Vol.41 (1962).

[198] H.J. Landau, H. Widom, *Eigenvalue distribution of time and frequency limiting*, J. Math. Analysis and Appl. Vol.77 (1980), 469–481.

[199] S. Lang, *Elliptic Functions*, (Second Edition), Graduate Texts in Mathematics, Vol.112, Springer-Verlag 1987.

[200] R.P. Langlands, *Automorphic representations, Shimura varieties, and motives. Ein Märchen.* in "Automorphic forms, representations and L-functions" Part 2, pp. 205–246, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., 1979.

[201] H. Laurent, *Sur la théorie des nombres premiers*. C.R. Acad. Sci. Paris, 126 (1898) 809–810.

[202] O. Lauscher, M. Reuter, *Asymptotic safety in quantum Einstein gravity: nonperturbative renormalizability and fractal spacetime structure*, hep-th/0511260.

[203] H. B. Lawson, M-L. Michelsohn *Spin geometry*, Princeton Mathematical Series, 38. Princeton University Press, Princeton, NJ, 1989.

[204] S. Lazzarini, T. Schucker, *A farewell to unimodularity*, Phys.Lett. B510 (2001) 277–284.

[205] E. Leichtnam, *An invitation to Deninger's work on arithmetic zeta functions*. in "Geometry, spectral theory, groups, and dynamics", 201–236, Contemp. Math., 387, Amer. Math. Soc., 2005.

[206] J.M. Lescure, *Triplets spectraux pour les variétés à singularité conique isolée.* Bull. Soc. Math. France 129 (2001), no. 4, 593–623.

[207] M. Levine, *Tate motives and the vanishing conjectures for algebraic K-theory*. in "Algebraic K-theory and algebraic topology" (Lake Louise, AB, 1991), 167–188, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 407, Kluwer Acad. Publ., 1993.

[208] M. Levine, *Mixed motives*, Math. Surveys and Monographs, Vol. 57, AMS, 1998.

[209] F. Lin, Y. Yang, *Gauged harmonic maps, Born-Infeld electromagnetism, and magnetic vortices*. Comm. Pure Appl. Math. 56 (2003), no. 11, 1631–1665.

[210] F. Lizzi, G. Mangano, G. Miele, G. Sparano, *Fermion Hilbert space and fermion doubling in the noncommutative geometry approach to gauge theories*, Phys. Rev. D 55 (1997) 6357–6366.

[211] J.L. Loday, *Cyclic homology.* Grundlehren der Mathematischen Wissenschaften, 301. Springer-Verlag, Berlin, 1998.

[212] D. Manchon, S. Paycha, *Renormalised Chen integrals for symbols on R^n and renormalised polyzeta functions*, preprint math.NT/0604562.

[213] Yu.I. Manin, *Correspondences, motifs and monoidal transformations*, Mat. Sb. (N.S.) 77 (119) 1968, 475–507.

[214] Yu.I. Manin, *Lectures on the K-functor in algebraic geometry.* Uspehi Mat. Nauk 24 1969 no. 5 (149), 3–86.

[215] Yu.I. Manin, *Theta functions, quantum tori and Heisenberg groups.* EuroConférence Moshé Flato 2000, Part III (Dijon). Lett. Math. Phys. 56 (2001), no. 3, 295–320.

[216] Yu.I. Manin, *Real Multiplication and noncommutative geometry*, in "The legacy of Niels Henrik Abel", pp.685–727, Springer, 2004.

[217] Yu.I. Manin, *Von Zahlen und Figuren*, in "Géomt́rie au XXe siècle. Histoire et horizons" (Ed. J. Kouneiher, D. Flament, Ph. Nabonnand, J.-J. Szczeciniarz), pp.24–44, Hermann, 2005,.

[218] Yu.I. Manin, *Lectures on zeta functions and motives (according to Deninger and Kurokawa).* Columbia University Number Theory Seminar (New York, 1992). Astérisque No. 228 (1995), 4, 121–163.

[219] Yu.I. Manin, M. Marcolli, *Continued fractions, modular symbols, and noncommutative geometry.* Selecta Math. (N.S.) 8 (2002), no. 3, 475–521.

[220] N.S. Manton, S.M. Nasir, *Volume of vortex moduli spaces.* Comm. Math. Phys. 199 (1999), no. 3, 591–604.

[221] M. Marcolli, *Arithmetic noncommutative geometry*, University Lectures Series, Vol.36, American Mathematical Society, 2005.

[222] J. Martinet, J.P. Ramis, *Elementary acceleration and multisummability, I*, Ann. Inst. Henri Poincaré, Vol.54 (1991) 331–401.

[223] J. Mather, *Commutators of diffeomorphisms. II*, Comment. Math. Helv. 50 (1975), 33–40.

[224] R. Meyer, *On a representation of the idele class group related to primes and zeros of L-functions.* Duke Math. J. Vol.127 (2005), N.3, 519–595.

[225] R. Meyer, R. Nest, *The Baum-Connes conjecture via localization of categories.* Lett. Math. Phys. 69 (2004), 237–263.

[226] J.S. Milne, *Canonical models of Shimura curves*, manuscript, 2003 (www.jmilne.org).

[227] J.S. Milne, *Fields and Galois theory*, manuscript, (www.jmilne.org).

[228] J.S. Milne, *Shimura varieties: the geometric side of the zeta function*, preprint (www.jmilne.org).

[229] J. Milnor, *Morse theory*, Annals of Mathematics Studies, No. 51, Princeton University Press, 1963.

[230] J. Milnor, *Introduction to algebraic K-theory*, Ann. Math. Studies, Vol.72, Princeton University Press, 1971.

[231] J. Milnor, J. Moore, *On the structure of Hopf algebras*, Ann. Math. (2) Vol.81 (1965) 211–264.

[232] J.A. Minahan, P. Ramond, R.C. Warner, *A comment on anomaly cancellation in the standard model*, Institute of Fundamental Theory Preprint UFIFT-HEP-89-16, June 1989, 6pp.

[233] R.N. Mohapatra, P.B. Pal, *Massive neutrinos in physics and astrophysics*, World Scientific, 2004.

[234] D. Mumford, *Tata Lectures on Theta. I*, Progress in Mathematics, Vol.28, Birkhäuser, Boston, MA, 1983.

[235] J.P. Murre, *Lecture on motives.* in "Transcendental aspects of algebraic cycles", 123–170, London Math. Soc. Lecture Note Ser., 313, Cambridge Univ. Press, Cambridge, 2004.

[236] S. Neshveyev, *Ergodicity of the action of the positive rationals on the group of finite adeles and the Bost-Connes phase transition theorem*, Proc. Amer. Math. Soc. Vol.130 (2002), N.10, 2999–3003.

[237] P. van Nieuwenhuizen, A. Waldron, *On Euclidean spinors and Wick rotations*, Phys. Lett. B 389 (1996) 29–36.

[238] J. Park, *Algebraic cycles and Connes periodicity*, preprint math.AG/0607272.

[239] S. Patterson, *An introduction to the theory of the Riemann Zeta-function*, Cambridge Univ. Press, 1988.

[240] F. Paugam, *Three examples of noncommutative boundaries of Shimura varieties*, in "Noncommutative Geometry and Number Theory" (Eds. C. Consani and M. Marcolli) pp.323–340, Vieweg Verlag, 2006.

[241] G.K. Pedersen, *C^*-algebras and their automorphism groups*, London Mathematical Society Monographs, 14. Academic Press, 1979.

[242] R. Percacci, *Renormalization group, systems of units and the hierarchy problem*, J. Phys. A: Math. Theor. 40 (2007) 4895–4913.

[243] A. Pilaftsis, *Gauge and scheme dependence of mixing renormalization*, Phys. Rev. D Vol.65 (2002) N.11 16pp.

[244] A. Polishchuk, *Noncommutative two-tori with real multiplication as noncommutative projective varieties*. J. Geom. Phys. 50 (2004), no. 1-4, 162–187.

[245] H.O. Pollak, D. Slepian, *Prolate spheroidal wave functions, Fourier analysis and uncertainty*, Bell Syst. Tech. Journal, Vol.40 (1961), 43–64.

[246] R.T. Powers, *Representations of uniformly hyperfinite algebras and their associated von Neumann rings*. Bull. Amer. Math. Soc. 73 1967 572–575.

[247] A. Pressley, G. Segal, *Loop groups*, Oxford Univ. Press, 1986.

[248] N. Ramachandran, *From Jacobians to one-motives: exposition of a conjecture of Deligne.* in "The arithmetic and geometry of algebraic cycles" (Banff, AB, 1998), 215–234.

[249] N. Ramachandran, *One-motives and a conjecture of Deligne*. J. Algebraic Geom. 13 (2004), no. 1, 29–80.

[250] P. Ramond, *Field theory: a modern primer*, Addison-Wesley, 1990.

[251] L. Reina, *TASI 2004 Lecture notes on Higgs boson physics*, hep-ph/0512377.

[252] A. Rennie, J.C. Varilly, *Reconstruction of manifolds in noncommutative geometry*, math.OA/0610418.

[253] M. Rieffel, *Morita equivalence for C^*-algebras and W^*-algebras*, J. Pure Appl. Algebra, 5 (1974), 51–96.

[254] B. Riemann, *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monat der Königl. Preuss. Akad. der Wissen. zu Berlin aus der Jahre 1859 (1860) 671–680. (English translation in M.H.Edwards "Riemann's zeta function", Dover 2001.)

[255] G. Ross, *Grand unified theories*, Frontiers in Physics Series, Vol.60, Benjamin, New York, (1985).

[256] C. Rovelli. *Statistical mechanics of gravity and the thermodynamical origin of time*. Classical Quantum Gravity, 10 (1993) N.8, 1549–1566.

[257] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, NewYork 1987.

[258] D. Ruelle, *Statistical mechanics. Rigorous results.* Reprint of the 1989 edition. World Scientific Publishing, Imperial College Press, London, 1999.

[259] N. Saavedra Rivano, *Catégories Tannakiennes.* Lecture Notes in Mathematics, Vol. 265. Springer-Verlag, 1972.

[260] A.J. Scholl, *Motives for modular forms.* Invent. Math. 100 (1990), no. 2, 419–430.

[261] E. Seiler, *Gauge theories as a problem of constructive Quantum Field Theory and Statistical Mechanics*, Lecture Notes in Physics, Vol.159, Springer, 1982.

[262] J.P. Serre, *Cours d'Arithmétique*, Presse Universitaire de France, 1977.

[263] J.P. Serre, *Motifs*, in "Journées Arithmétiques, 1989" (Luminy, 1989). Astérisque No. 198-200 (1991), 11, 333–349 (1992).

[264] J.P. Serre, *Propriétés conjecturales des groupes de Galois motiviques et des representations l-adiques*, Proc. Symp. Pure Math. Vol. 55 (1994) Part I, 377–400.

[265] J. P. Serre, *Facteurs locaux des fonctions zêta des variétés algébriques (définitions et conjectures).* Sém. Delange-Pisot-Poitou, exp. 19, 1969/70.

[266] J.P. Serre, *Analogues kählériens de certaines conjectures de Weil*, Ann. of Math. (2) 71 (1960) 392–394.

[267] M. Sher, *Electroweak Higgs potential and vacuum stability*, Phys. Rep. Vol.179 (1989) N.5-6, 273–418.

[268] G. Shimura, *Arithmetic theory of automorphic functions*, Iwanami Shoten and Princeton 1971.

[269] T. Shiota, *Characterization of Jacobian varieties in terms of soliton equations.* Invent. Math. 83 (1986), no. 2, 333–382.

[270] L. Sibner, R. Sibner, Y. Yang, *Abelian gauge theory on Riemann surfaces and new topological invariants* R. Soc. Lond. Proc. Ser. A, Math. Phys. Eng. Sci. Vol.456 (2000), N.1995, 593–613.

[271] R. Sips, *Representation asymptotique des fonctions de Mathieu et des fonctions d'onde spheroidales*, Trans. Am. Math. Soc. Vol.66 (1949), 93–134.

[272] D. Slepian, *Some asymptotic expansions for prolate spheroidal wave functions*, J. Math. Phys. Vol.44 (1965), 99–140.

[273] J. Steenbrink, *A summary of mixed Hodge theory*, in "Motives" vol.I, Proc. Symp. Pure Math. Vol.55, AMS 1991, pp.31–42.

[274] K.S. Stelle, *Renormalization of higher-derivative quantum gravity*, Physical Review D, Vol.16 (1977) N.4, 953–969.

[275] C.A. Stephan, *Almost-Commutative Geometry, massive Neutrinos and the Orientability Axiom in KO-Dimension 6*, preprint hep-th/0610097.

[276] W. van Suijlekom, *The Hopf algebra of Feynman graphs in QED*, preprint hep-th/0602126.

[277] W. van Suijlekom, *Renormalization of gauge fields: A Hopf algebra approach*, preprint hep-th/0610137.

[278] H. P. F. Swinnerton-Dyer, *Applications of Algebraic Geometry to Number Theory*, Proceedings Symposia in Pure Math. Vol. XX (1969) 1–52.

[279] M. Takesaki, *Tomita's theory of modular Hilbert algebras and its applications.* Lecture Notes in Math., 128, Springer, 1970.

[280] M. Takesaki, *Duality for crossed products and the structure of von Neumann algebras of type III*, Acta Math. Vol.131 (1973), 249–310.

[281] J. Tate, *Fourier analysis in number fields and Hecke's zeta-function*, Ph.D. Thesis, Princeton, 1950. Reprinted in J.W.S. Cassels and A. Frölich (Eds.) "Algebraic Number Theory", Academic Press, 1967.

[282] J. Tate, *Number theoretic background*, in "Automorphic forms, representations, and L-functions", Proc. Symp. Pure Math. Vol.33, Part 2, 1979, pp.3–26.

[283] M. van der Put, *Differential Galois Theory, Universal Rings and Universal groups*, in "Differential Algebra and Related topics", Editors: Li Guo, Phyllis Cassidy, William F. Keigher, William Sitt. World Scientific, 2002.

[284] M. van der Put, M. Singer, *Galois theory of linear differential equations*, Springer 2002.

[285] M. Veltman, *Diagrammatica: the path to Feynman diagrams*, Cambridge Univ. Press, 1994.

[286] V. Voevodsky, *Motivic cohomology groups are isomorphic to higher Chow groups in any characteristic.* Int. Math. Res. Not. 2002, no. 7, 351–355.

[287] V. Voevodsky, *Triangulated categories of motives over a field* in "Cycles, transfer and motivic homology theories, pp. 188–238, Annals of Mathematical Studies, Vol. 143, Princeton, 2000.

[288] J. von Neumann, *On rings of operators. Reduction theory.* Ann. of Math. (2) 50 (1949) 401–485.

[289] R. M. Wald, *General Relativity*, The University of Chicago Press, Chicago and London (1984).

[290] W.C. Waterhouse, *Introduction to affine group schemes*. Graduate Texts in Mathematics, 66. Springer, 1979.

[291] N.E. Wegge-Olsen, *K-Theory and C^*-Algebras: A Friendly Approach*. Oxford, England: Oxford University Press, 1993.

[292] A. Weil, *Letter to Artin*, Collected Papers, Vol.I (1980) 280–298.

[293] A. Weil, *Basic Number Theory*, Reprint of the second (1973) edition. Classics in Mathematics. Springer-Verlag, 1995.

[294] A. Weil, *Sur les formules explicites de la théorie des nombres premiers*, Izv. Mat. Nauk., (Ser. Mat.) Vol.36 (1972) 3–18. (in Oeuvres compltes, Vol. 2, 48–62.)

[295] A. Weil, *Sur certains groupes d'opérateurs unitaires*, Acta Math. , 111, (1964).

[296] A. Weil, *Fonction zeta et distributions*, Séminaire Bourbaki, Vol. 9 (1966), Exp. No. 312, 523–531.

[297] A. Weil, *Elliptic functions according to Eisenstein and Kronecker*, Springer, 1976.

[298] S. Weinberg, *Gravitation and Cosmology*, John Wiley and sons (1972).

[299] E. Witten, *Noncommutative geometry and string field theory*. Nuclear Phys. B 268 (1986), 253–294.

[300] E. Witten, *An SU(2) anomaly*, Phys. Lett. B, Vol.117 (1982) 324–328.

[301] E. Witten, $2+1$-*dimensional gravity as an exactly soluble model,* Nucl. Phys. B311 (1988) 46–78.

[302] E. Witten, *Three-dimensional gravity reconsidered*, arXiv:0706.3359

[303] M. Wodzicki. *Noncommutative residue, Part I. Fundamentals*, Lecture Notes Math., 1289, Springer, Berlin, (1987).

[304] J. Wood, *Unbounded multipliers on commutative Banach algebras*, Pacific J. Math. Vol.85 (1979) no. 2, 479–484.

[305] W. Zimmermann, *Convergence of Bogoliubov's method of renormalization in momentum space*, Comm. Math. Phys. 15, (1969), 208–234.