LECTURES ON ANALYTIC NUMBER THEORY

J. R. QUINE

Contents

1. What is Analytic Number Theory?

From D. J. Newman [5]: *The most intriguing thing about Analytic Number Theory (the use of analysis, or function theory, in number theory) is its very existence! How could one use properties of continuous valued functions to determine properties of those most discrete items, the integers. Analytic functions? What has differentiability got to do with counting? The astonishment mounts further when we learn that the complex zeros of a certain analytic function are the basic tools in the investigation of the primes. The answer to all this bewilderment is given by two words, generating functions* [5]

1.1. **Generating functions.** The simplest kind of generating function is a polynomial or power series $\sum a_k z^k$. Assume in general that the sum of a power series is from $k = 0$ to ∞. For analytic number theory the coefficients a_k should be functions of integers and the series, if it converges, a function of z that can be studied by calculus or by analytic function theory.

A simple example of how analysis can be used to get a number theory result is found by letting $a_k = k$ be the sequence of integers from $k = 1$ to $k = n$. A generating function using this sequence of integers as coefficients is

$$\begin{aligned}\sum_{k=1}^{n} kz^{k-1} &= \frac{d}{dz}\sum_{k=0}^{n} z^k = \frac{d}{dz}\frac{z^{n+1}-1}{z-1} \\ &= \frac{nz^{n+1}-(n+1)z^n+1}{(z-1)^2}\end{aligned}$$

Now taking the limit as $z \to 1$ and using l'Hospital's rule on the function on the right, get

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

the well known formula for the sum of the first n integers. Thus we have obtained a formula for integer sums using calculus.

Exercise: Using a similar technique, show that

$$\sum_{k=1}^{n} k^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

Another type of generating function different from a power series is a Dirichlet series, $\sum a_n n^{-s}$. Assume in general the sum of a Dirichlet series is from $n = 1$ to ∞. The simplest such series is the Riemann zeta function $\sum n^{-s}$ where all of the coefficients a_n are equal to 1. The sum converges for $s > 1$. We can also take finite Dirichlet sums, for example, $\sum_{n=1}^{N} n^{-s}$. When $s = -2$ or $s = -1$ we get the sums $\sum_{k=1}^{n} k$ and $\sum_{k=1}^{n} k^2$. For $-s$ a non-negative integer, the sums are called *Faulhaber sums*.

We can form a generating function using Faulhaber sums for coefficients of a power series. For μ a non-negative integer and N a positive integers let $S(\mu, N) = \sum_{n=1}^{N} n^{\mu}$. Now write a generating power series using these sums in the coefficients. Interchanging the order of summation, and using the power series for the exponential function get

$$\sum_{\mu=0}^{\infty} \frac{1}{\mu!} S(\mu, N) z^{\mu} = \sum_{n=1}^{N} e^{nz} = \frac{e^{(N+1)z} - e^z}{e^z - 1}.$$

Calculating the coefficients of the series on the right gives of way of finding formulas for the Faulhaber sums. See [3].

1.2. **Operations on series.** In using these generating functions, we use rules for adding, multiplying, and dividing to create other functions. We can do the formal computation without worrying about convergence.

1.2.1. *Addition.* Both Dirichlet and power series can be added term by term.

We can form an infinite sum of power series $\sum_{k=0}^{\infty} h_k$ if $h_k = \mathrm{O}(z^k)$, that is, the first non zero coefficient of h_k is the k^{th} one. We need this so the sums of the coefficients can be collected as finite sums and we get another well-defined power series.

Under similar conditions we can add an infinite number of Dirichlet series.

1.2.2. *Multiplication.* For both Dirichlet and power series we can multiply term by term and collect like terms to get another series of the same type. Multiplication of power series:

$$\begin{aligned}\sum_m a_m z^m \sum_n b_n z^n &= \sum_{m,n} a_m b_n z^{m+n} \\ &= \sum_k \left(\sum_{m+n=k} a_m b_n \right) z^k\end{aligned}$$

where sums are from 0 to ∞. The coefficient

$$c_n = \sum_{m+n=k} a_m b_n$$

of the product can be thought of as the sum along the diagonal line $m + n = k$ of the product of the coefficients.

Multiplication of Dirichlet series:

$$\sum_m a_m m^{-s} \sum_n b_n n^{-s} = \sum_{m,n} a_m b_n (mn)^{-s}$$
$$= \sum_k \left(\sum_{mn=k} a_m b_n \right) k^{-s}$$

where sums are from 1 to ∞. The coefficient

$$c_n = \sum_{mn=k} a_m b_n$$

of the product can be thought of as the sum along the hyperbola $mn = k$ of the product $a_m b_n$ of the coefficients.

In some cases it makes sense to take the infinite product of power series or Dirichlet series.

1.2.3. *Inverse.* For taking the inverse of a power series assume that $a_0 = 1$ and the series is $1 + h$ where $h = \sum_{k=1}^{\infty} a_k z^k$. Then use the rule for a geometric series

$$\frac{1}{1+h} = \sum_{k=0}^{\infty} (-1)^k h^k$$

A similar method works for finding the inverse of a Dirichlet series.

As an example of inverting a power series using the above method, consider the series

$$\frac{e^z - 1}{z} = 1 + \sum_{k=1}^{\infty} \frac{z^k}{(k+1)!}.$$
$$= 1 + \left(\frac{z}{2} + \frac{z^2}{6} + \cdots \right)$$

Then

$$\frac{z}{e^z - 1} = 1 - \left(\frac{z}{2} + \frac{z^2}{6} + \cdots \right) + \left(\frac{z}{2} + \frac{z^2}{6} + \cdots \right)^2 + \cdots$$
$$= 1 - \frac{1}{2} z + \frac{1}{12} z^2 - \cdots$$

If we write

$$\frac{z}{e^z - 1} = \sum \frac{B_k}{k!} z^k$$

then the above method is a way of generating the numbers B_k, called the Bernoulli numbers. We see that $B_0 = 1$, $B_1 = -1/2$ and $B_2 = 1/6$.

As an example of inverting Dirichlet series, consider the series $\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = 1 + h$ where $h = \sum_{n=2}^{\infty} n^{-s}$. Then

$$\frac{1}{\zeta(s)} = \sum_{k=0}^{\infty} (-1)^k h^k$$

So

$$\begin{aligned}\frac{1}{\zeta(s)} &= 1 - \left(2^{-s} + 3^{-s} + 4^{-s} + 5^{-s} + 6^{-s} + \cdots\right) \\ &\qquad + \left(2^{-s} + 3^{-s} + 4^{-s} + \cdots\right)^2 - \cdots \\ &= 1 - 2^{-s} - 3^{-s} - 5^{-s} + 6^{-s} + \cdots\end{aligned} \tag{1}$$

If we write

$$\frac{1}{\zeta(s)} = \sum_n \mu(n) n^{-s}$$

then the above method is a way of generating the numbers $\mu(n)$. It looks like all the numbers will be 1 or -1 or 0. There is a simple expression for these numbers in terms of primes, but it is not clear how to find it using the above method.

Exercise: Generate some more of the sequence a_n and try to guess a simple formula for a_n .

Addition, multiplication and inversion give ways to create new series from given ones. Also the exponential and logarithm can be taken, as we will see. The coefficients of the resulting series may be of number theoretic interest.

1.3. **Some interesting series.** We give two example of interesting generating functions, both the subject of much study. Both can be represented as an infinite product, a fact which gives them their number theoretic interest.

The first is

$$\frac{1}{(1-z)(1-z^2)(1-z^3)\cdots} = \sum_{n=0}^{\infty} p(n) z^n \tag{2}$$

where $p(k)$ is the partition function, the number of ways of representing the integer n as the sum of positive integers, not counting the order in which the sum is written (Wikipedia reference). For example

$$\begin{aligned}5 &= 5 \\ &= 4+1 \\ &= 3+1+1 \\ &= 3+2 \\ &= 2+2 \\ &= 2+1+1+1 \\ &= 1+1+1+1+1\end{aligned}$$

so $p(5) = 7$. Equation (2) can be checked by writing each factor $1/(1-x^k)$ as a geometric series $\displaystyle\sum_{j=0}^{\infty} x^{jk}$.

The second is the Riemann zeta function and it's Euler factorization,

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_p (1-p^{-s})^{-1} \tag{3}$$

where the product on the right is the product over all primes. This factorization is equivalent to the statement that every integer has a unique factorization into primes. Equation (3) can be checked formally by writing each factor $1/(1-p^{-s})$

as a geometric series $\sum_{j=0}^{\infty} p^{-sj}$. This factorization explains why we got only $1, -1, 0$ as coefficients in the series (1) for $1/\zeta$.

Analysis of the function (2) yields the

Asymptotic formula for the partition function

$$p(n) \sim \frac{1}{4\sqrt{3}n} \exp\left(\pi\sqrt{\frac{2n}{3}}\right) \text{ as } n \to \infty.$$

There is proof of this in Chapter II of [5].

Analysis of the function (3) yields the

Prime Number Theorem. *If $\pi(x)$ is the number of primes $\leq x$ then*

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \to \infty.$$

There is a short proof of this in the paper [8]. A longer proof is given in [6] Chapter 7. A good history of the prime number theorem is found in [2].

Dirichlet L functions are a generalization of the zeta function. Analysis of these shows

Dirichlet's Theorem *If q and ℓ are relatively prime positive integers, then there are infinitely many primes of the form $\ell + kq$ with $k \in \mathbb{Z}$.*

The proof of this is covered in [7] Chapter 6, and we will discuss this in the next sections.

2. The Zeta Function

2.1. **Some elementary number theory.** The following elementary facts are proved in Chapter 6 of [7].

(1) **Euclid's Algorithm.** *For any integers a and b with $b > 0$, there exist unique integers q and r with $0 \leq r < b$ such that*

$$a = qb + r.$$

(2) *If $\gcd(a, b) = d$ then there exists integers x and y such that*

$$xa + yb = d.$$

(3) *Two positive integers are relatively prime if and only if there exist integers x and y such that*

$$xa + yb = 1.$$

(4) *If a and c are relatively prime and c divides ab then c divides b. In particular, if p is a prime that does not divide a, and p divides ab, then p divides b.*

(5) *If p is prime and p divides the product $a_1 \cdots a_n$, then p divides a_i for some i.*

(6) *n is relatively prime to q if and only if there is m such that $mn = 1 \bmod q$.*

(7) **Fundamental theorem of arithmetic.** *Every positive integer greater than 1 can be factored uniquely into the product of primes.*

2.2. **The infinitude of primes.** Euclid provided a simple argument that there are an infinite number of primes. A simple variation of this argument shows that

There are an infinite number of primes of the form $4k + 3$.

Proof. Proof by contradiction. Suppose $p_1 = 7, p_2 = 11, \ldots, p_n$ are all primes $\equiv 3 \bmod 4$, except for 3. Let $N = 4p_1 \cdots p_n + 3$. The product of integers $\equiv 1 \bmod 4$ is and integer $\equiv 1 \bmod 4$, so if all prime factors of N are $\equiv 1 \bmod 4$ then N would be $\equiv 1 \bmod 4$. Since $N \equiv 3 \bmod 4$ there must be a prime factor $\equiv 3 \bmod 4$. This factor cannot be $p_1, \ldots, p_n$ or 3, since these are not factors of N. Therefore there must be another prime $\equiv 3 \bmod 4$ other than the ones listed. □

The same proof does not work for primes of the form $4k + 1$, so the question arises whether there are an infinite number of primes of the form $4k + 1$. Here we need to use some analysis of the zeta function and the Euler product and to define an L function.

2.3. **Infinite products.** For a sequence A_n of complex numbers, define the infinite product by

$$\prod_{n=1}^{\infty} A_n = \lim_{N\to\infty} \prod_{n=1}^{N} A_n.$$

For a sequence of non-zero numbers you can show the product converges by showing the sum of the logarithms converges. Using this method we find that

If $A_n = 1 + a_n$ *and* $\sum |a_n|$ *converges, the the product* $\prod_n A_n$ *converges, and this product is zero if and only if one of the factors* A_n *is zero. Also if* $a_n \neq 1$ *for all* n*, then* $\prod_n 1/(1 - a_n)$ *converges.*

The proof relies on the following inequalities which will be used often. If $|z| < 1/2$

$$\begin{aligned} |\log(1+z) - z| &\le |z|^2 \\ |\log(1+z)| &\le 2|z|. \end{aligned} \tag{4}$$

For complex numbers we use the principal branch of the log.

2.4. **The zeta function and Euler product.** For $s > 1$ the zeta function is defined by

$$\zeta(s) = \sum_{n=1} n^{-s}.$$

By the integral test, the series converges. By the comparison test it converges uniformly for $s \ge s_0 > 1$, and therefore is a continuous function of s.

The zeta function is useful in the study of primes because it can be factored into an infinite product, the Euler product, using primes.

For every $s > 1$ *we have*

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}. \tag{5}$$

where the product is taken over all primes.

Taking the log of equation (5) we find that

$$\log \zeta(s) = -\sum_p \log(1 - p^{-s}) = \sum_p p^{-s} + \mathrm{O}(1)$$

where the O(1) term is $\leq \sum_{n=1}^{\infty} n^{-2}$ independent of s. Letting $s \to 1^+$ get

$$\sum_p 1/p = \infty, \tag{6}$$

proving in a different way from Euler that there are an infinite number of primes. Although this is not the easiest way to prove there are an infinite number of primes, the method generalizes to prove much more interesting results.

2.5. **Infinitude of primes of the form $4k + 1$.** Using (6) and a variation of the zeta function, called a L function, it can be shown that there are an infinite number of primes of the form $4k + 1$.

We first define a function χ on $\mathbb{Z}$ by

$$\chi(n) = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{if } n \equiv 1 \bmod 4 \\ -1 & \text{if } n \equiv 3 \bmod 4 \end{cases} \tag{7}$$

and check directly that χ has the multiplicative property $\chi(mn) = \chi(m)\chi(n)$ on all of $\mathbb{Z}$. The function χ is called a Dirichlet character.

Let

$$L(s, \chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s} = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots . \tag{8}$$

for $s > 1$. The multiplicative property of the character implies that if an integer n factors into primes as $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ then $\chi(n) n^{-s} = (\chi(p_1) p_1^{-s})^{\alpha_1} \cdots (\chi(p_k) p_k^{-s})^{\alpha_k}$. Using this fact, for $s > 1$ an Euler product can be written for L,

$$L(s, \chi) = \prod_p (1 - \chi(p) p^{-s})^{-1} \tag{9}$$

with a similar proof as for $\zeta(s)$. A detailed proof will be given later.

The function L looks similar the ζ function, but because the series (8) is alternating, it has a finite value at $s = 1$ by the alternating series test. We can see that it converges to a value between 1 and 2/3. A more detailed analysis shows it converges to $\pi/4$, and that

$$\lim_{s \to 1^+} L(s, \chi) = L(1, \chi),$$

a fact that will be show later using summation by parts to show that the function L is continuous for $0 < s$.

Now arguing as with the zeta function,

$$\log L(s, \chi) = \sum_p \chi(p) p^{-s} + \mathrm{O}(1).$$

Conclude that $\sum_p \chi(p) p^{-s}$ remains bounded as $s \to 1^+$.

Write for $s > 1$,

$$\sum_p \chi(p) p^{-s} = \sum_{p \equiv 1} p^{-s} - \sum_{p \equiv 3} p^{-s}$$

and

$$\sum_p p^{-s} = 2^{-s} + \sum_{p \equiv 1} p^{-s} + \sum_{p \equiv 3} p^{-s}$$

where all congruences are mod 4. The sum is

$$2^{-s} + 2\sum_{p\equiv 1} p^{-s}.$$

and so

$$\sum_{p\equiv 1} p^{-s} \to \infty \text{ as } s \to 1^+$$

and consequently

$$\sum_{p\equiv 1} p^{-1} = \infty.$$

This shows there are an infinite number of primes $\equiv 1$ mod 4.

3. Dirichlet characters and L functions

The goal is to prove Dirichlet's theorem.

Dirichlet's theorem. *If q and ℓ are relatively prime positive integers, the there are infinitely many primes of the form $\ell + kq$ with $k \in \mathbb{Z}$.*

We have already proved the theorem for $q = 4$, $\ell = 1, 3$. The proof for $\ell = 1$ uses the zeta function and illustrates the proof for arbitrary q. It also is an introduction to L functions. Generalizing the L function requires a study of Dirichlet characters.

3.1. Dirichlet characters.

Recall that $\mathbb{Z}(q)$ is the group of integers mod q. Define $\mathbb{Z}(q)^*$ to be the group of units of $\mathbb{Z}(q)$ under multiplication. The set of units is the set of ℓ such that $\gcd(\ell, q) = 1$. The number of elements of $\mathbb{Z}(q)^*$ is denoted $\varphi(q)$ which is also called *Euler's totient function*.

A *character mod q* is a map $\chi : \mathbb{Z}(q)^* \to S^1 \subset \mathbb{C}$ such that $\chi(mn) = \chi(m)\chi(n)$ for all m, n. Note that S^1 is the unit circle in the complex plane. We will also call these *Dirichlet characters*.

We may also think of $\chi(n)$ as defined for any integer relatively prime to q by identifying the integer n with it's equivalence class in $\mathbb{Z}(q)^*$. The character can then be extended to all of $\mathbb{Z}$ by setting $\chi(n) = 0$ if n is not relatively prime to q. Use the same symbol χ for the character on $\mathbb{Z}(q)^*$ and the character extended to $\mathbb{Z}$. It is clear that $\chi(mn) = \chi(m)\chi(n)$ (the multiplicative property) also holds for the character extended to all of $\mathbb{Z}$.

For any q we can define the *trivial character* by setting $\chi(m)$ to be 1 for all $m \in \mathbb{Z}(q)^*$. The trivial character is denoted χ_0.

The character χ above we used for $q = 4$ is defined on $\mathbb{Z}(4)^* = \{1, 3\}$ by setting $\chi(1) = 1$ and $\chi(3) = -1$.

3.2. Construction of Dirichlet characters.

In this section we show that we can define Dirichlet characters for all integers q. We also investigate properties of the set of characters.

To illustrate the method we will construct characters for a few values of q. In the previous section we constructed a character for $q = 4$. There is also the trivial character giving 2 in total.

Lets try a similar method for $q = 5$. The elements of $\mathbb{Z}(q)^*$ are 1,2,3,4. Note that $2^2 = 4$ and $2^3 = 3$, so the elements of $\mathbb{Z}(q)^*$ are 2^j for $j = 0, 1, 2, 3$. Note that the values of 2^j are the same for any j in the same residue class mod 4, since $2^4 \equiv 1$. This shows that $\mathbb{Z}(5)^*$ is isomorphic to the cyclic group $\mathbb{Z}(4)$.

If z is any fourth root of unity we define a character by setting $\chi(2^j) = z^j$. This is well defined since the right hand side has the same value for values of j in the same residue class mod 4. It is clear that χ is a character since if $m = 2^j$ and $n = 2^k$ then

$$\chi(m)\chi(n) = \chi(2^j)\chi(2^k) = z^j z^k = z^{j+k} = \chi(mn).$$

We get a character for z equal to each fourth root of unity, $1, i, -1, -i$. When $z = 1$ we get the trivial character.

We can summarize the list of characters in table 1. More conveniently it can be

TABLE 1. Dirichlet characters for $\mathbb{Z}(5)^*$

$\chi\backslash n$	1	2	3	4
$\chi_0(n)$	1	1	1	1
$\chi_1(n)$	1	i	$-i$	-1
$\chi_2(n)$	1	-1	-1	1
$\chi_3(n)$	1	$-i$	i	-1

summarized as a 4×4 matrix,

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -i & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -i & i & -1 \end{pmatrix} \tag{10}$$

Notice that A has the property that $A^*A = 4I$ so that $\frac{1}{2}A$ is unitary. We will see that this generalizes to characters of $\mathbb{Z}(q)^*$ arbitrary q.

For $q = 8$, the situation is different since the group $\mathbb{Z}(8)^*$ is not cyclic. The elements of $\mathbb{Z}(8)^*$ are $1, 3, 5, 7$. We have $3^2 = 5^2 = 7^2 = 1$ so the elements are not powers of any one element. However $7 = 3 \cdot 5$ so every element of $\mathbb{Z}(8)^*$ can be written as $3^j 5^k$ for $j = 0, 1$ and $k = 0, 1$ and the expression is the same if j or k are change by a multiple of 2. This shows that $\mathbb{Z}(8)^*$ is isomorphic to the direct sum of cyclic groups $\mathbb{Z}(2) \oplus \mathbb{Z}(2)$.

Now taking z_1 and z_2 to be square roots of 1, that is, 1 or -1, we can get a character by defining $\chi(3^j 5^k) = z_1^j z_2^k$. As before it is easy to see that this has the multiplicative property. The four characters are given by the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \tag{11}$$

where the column correspond to the elements 1, 3, 5, 7 of $\mathbb{Z}(8)^*$, and the rows correspond to the four different characters.

The examples above illustrate a method for proving the following theorem.

Existence and orthogonality of characters *For every positive integer $q \geq 2$ there are exactly $\varphi(q)$ characters of $\mathbb{Z}(q)^*$. The set of characters is a group under multiplication. This group is isomorphic to $\mathbb{Z}(q)^*$. When the characters are written in the form of a matrix A,*

$$AA^* = \varphi(q)I.$$

Proof. We need to show that the construction of characters illustrated above works for any q. For this we need the fact that every finite abelian group can be written as the direct product of cyclic groups. We will not prove this fact. In the case of the group $\mathbb{Z}(q)^*$ it means that there are elements $a_1, \dots, a_k$ of order $N_1, \cdots N_\ell$ respectively such that every element of $\mathbb{Z}(q)^*$ can be written uniquely as $a_1^{j_1} \cdots a_k^{j_k}$ for $0 \le j_\ell \le N_\ell - 1$, where $N_1 \cdots N_\ell = \varphi(q)$. We saw this in the case $q = 8$ above with $a_1 = 3, a_2 = 5$, $N_1 = N_2 = 2$.

Now letting $\omega_\ell = e^{2\pi i/N_l}$ and defining

$$\chi(a_1^{j_1} \cdots a_k^{j_k}) = \omega_1^{n_1 j_1} \cdots \omega_k^{n_k j_k} \tag{12}$$

gives a character for each choice of integers $0 \le n_\ell \le N_\ell - 1$. This gives $\varphi(q)$ characters.

The product of characters is clearly a character. Associate the character defined in (12) with the element $a_1^{n_1} \cdots a_k^{n_k}$ in $\mathbb{Z}(q)^*$. This is easily seen to be a group isomorphism.

Next we show that $AA^* = \varphi(q)I$. This is equivalent to

$$\sum_{n \in \mathbb{Z}(q)^*} \overline{\chi_1(n)} \chi_2(n) = \begin{cases} 0 & \text{if } \chi_1 \ne \chi_2 \\ \varphi(q) & \text{if } \chi_1 = \chi_2. \end{cases} \tag{13}$$

If $\chi_1 \ne \chi_2$ then there is some m such that $\chi_1(m) \ne \chi_2(m)$. Now

$$\begin{aligned} \sum_{n \in \mathbb{Z}(q)^*} \overline{\chi_1(n)} \chi_2(n) &= \sum_{n \in \mathbb{Z}(q)^*} \overline{\chi_1(mn)} \chi_2(mn) \\ &= \overline{\chi_1(m)} \chi_2(m) \sum_{n \in \mathbb{Z}(q)^*} \overline{\chi_1(n)} \chi_2(n). \end{aligned}$$

The first equality is a change in the order of summation, and the second is by the multiplicative property of characters. Since $\overline{\chi_1(m)}\chi_2(m) \ne 1$, it follows that $\sum_n \overline{\chi_1(n)} \chi_2(n) = 0$. If $\chi_1 = \chi_2$, all the terms in the sum are 1 and so the sum is the number of elements in $\mathbb{Z}(q)^*$, $\varphi(q)$.

□

We can think of a character as a vector in $\mathbb{C}^{\varphi(q)}$, a row of A. The above expression then becomes $\langle \chi_1, \chi_2 \rangle$, the Hermitian inner product. The statement $\langle \chi_1, \chi_2 \rangle = 0$ means that the rows of A are orthogonal.

Since $A^*/\sqrt{\varphi(q)}$ is the inverse of $A/\sqrt{\varphi(q)}$, we also have that $A^*A = \varphi(q)I$ and so the columns of A are orthogonal. This is written as

$$\sum_{\chi} \overline{\chi(\ell)} \chi_(n) = \begin{cases} 0 & \text{if } \ell \ne n \\ \varphi(q) & \text{if } \ell = n. \end{cases} \tag{14}$$

In the proof we used the fact that every finite abelian group can be written as the direct product of cyclic groups. Alternately we can use the fact that a commuting family of unitary transformation on a finite dimensional space are simultaneously diagonalizable. See [7] page 233.

The fact that the group of characters and $\mathbb{Z}(q)^*$ are isomorphic is used by Newman in his proof in [5] of the non vanishing of $L(1, \chi)$ for non trivial characters.

3.3. **Euler product for L functions.** Similar to the proof of the Euler product for the zeta function we can show that for $s > 1$,

$$L(s,\chi) = \prod_p (1 - \chi(p)p^{-s})^{-1}. \tag{15}$$

The proof uses the fundamental theorem of arithmetic. The basic idea is to make sense of the statement

$$\lim_{N\to\infty} \sum_{n\le N} \chi(n)n^{-s} = \lim_{M,N\to\infty} \prod_{p\le N} \sum_{k_p=0}^{M} \chi(p)p^{-k_p s}.$$

The finite sums on either side have approximately the same terms by the fundamental theorem of arithmetic and by the multiplicative property of characters. See [7] p. 260.

Denoting the sum on the left by S_N and the one on the right by $\prod_{N,M}$, the key step is to estimate $|S_N - \prod_{N,M}|$. Note that if $p^\alpha | n$ and $n \le N$ then $p \le N$ and $\alpha \le \log N/\log p \le \log N/\log 2$. It follows from the fundamental theorem of arithmetic that if $M > \log N/\log 2$, the product $\prod_{N,M}$ multiplied out into a sum contains all the terms of S_N exactly once. We can then estimate

$$\left| S_N - \prod\nolimits_{N,M} \right| \le \left| \sum_{n>N} \chi(n)n^{-s} \right| \le \sum_{n>N} n^{-s},$$

and the sum on the right goes to zero as $N \to \infty$ for $s > 1$.

3.4. **Outline of proof of Dirichlet's theorem.** The proof of Dirichlet's theorem for arbitrary q follows the method used for $q = 4$. Consider the $\varphi(q)$ characters χ on $\mathbb{Z}^*(q)$. By the orthogonality of characters, for ℓ and m in $\mathbb{Z}^*(q)$,

$$\varphi(q)\delta_\ell(m) = \sum_\chi \overline{\chi(\ell)}\chi(m) \tag{16}$$

where

$$\delta_\ell(m) = \begin{cases} 1 & \text{if } \ell = m \\ 0 & \text{if } \ell \ne m. \end{cases}$$

If the characters are extended to $m \in \mathbb{Z}$, equation (16) also holds with $\delta_\ell(m)$ defined by

$$\delta_\ell(m) = \begin{cases} 1 & \text{if } m \equiv \ell \bmod q \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Write (16) as

$$\varphi(q)\delta_\ell(m) = \chi_0(m) + \sum_{\chi\ne\chi_0} \overline{\chi(\ell)}\chi(m) \tag{18}$$

where χ_0 is the trivial character. Multiply both sides of (18) by m^{-s} and sum over all $m = p$ where p is a prime to get

$$\varphi(q)\sum_{p\equiv\ell} p^{-s} = \sum_{p\nmid q} p^{-s} + \sum_{\chi\ne\chi_0} \overline{\chi(\ell)} \sum_p \chi(p)p^{-s}. \tag{19}$$

Now

$$\sum_{p\nmid q} p^{-s} = \sum_p p^{-s} - \sum_{p|q} p^{-s}$$

and the sum on the right is a finite sum. Since it was shown that

$$\sum_p p^{-s} \to \infty \text{ as } s \to 1^+$$

it follows that

$$\sum_{p \nmid q} p^{-s} \to \infty \text{ as } s \to 1^+.$$

If $\sum_p \chi(p)p^{-s}$ is bounded as $s \to 1^+$ for every non trivial character χ, then by (19)

$$\sum_{p \equiv \ell} p^{-1} = \infty,$$

so there are an infinite number of prime $p \equiv \ell \bmod q$ and this proves Dirichlet's theorem.

As before

$$\log L(s, \chi) = \sum_p \chi(p)p^{-s} + \mathrm{O}(1)$$

as $s \to 1^+$, so to complete the proof of Dirichlet's theorem it is left to show that $L(1, \chi)$ is finite and non-zero for every non trivial character χ, and $\lim_{s \to 1^+} \log L(s, \chi) = L(1, \chi)$. This will be proved in following sections.

4. Analytic tools

Before continuing with L functions and the proof of Dirichlet's theorem, review a few tools from analysis.

4.1. **Summation by parts.**

Summation by parts is analogous to integration by parts.

Let a_j and b_j be sequences and let $A_n = \sum_{j=1}^n a_j$, then for integers $M > N$,

$$\sum_{n=M}^{N} a_n b_n = \sum_{n=M}^{N-1} A_n(b_n - b_{n+1}) + A_N b_N - A_{M-1} b_M. \tag{20}$$

Proof Write $\displaystyle\sum_{n=M}^{N} a_n b_n = \sum_{n=M}^{N} (A_n - A_{n-1}) b_n$ and break the sum into two parts. Change the index of summation from n to $n+1$ in the second part. □

The formula for summation by parts can be written in a different way if f is a differentiable function and $b_n = f(n)$. In this case write

$$b_{n+1} - b_n = \int_n^{n+1} f'(x)\,dx,$$

and $A(x) = \sum_{n \le x} a_n$. The sum on the right of (20) becomes

$$-\int_M^N A(x) f'(x)\,dx.$$

The sum on the left is written

$$\int_M^N f(x) dA(x).$$

This is really a *Stieltjes integral*, but can be thought of simply as a sum. This gives a way to remember summation by parts, because it can be written in the same way as integration by parts.

Summation by parts, integral form. *Let a_j be a sequence and f a differentiable function. Let $A(x) = \sum_{j \le x} a_j$, then for integers $M > N$,*

$$\int_M^N f(x) dA(x) = A(x)f(x)\big|_{M^-}^N - \int_M^N A(x)f'(x)\,dx. \tag{21}$$

The M^- indicates the limit from the left.

Summation by parts can be applied to Dirichlet series by setting $f(x) = x^{-s}$ above. In this case the summation by parts formula becomes

$$\sum_{n=M}^N a_n n^{-s} = s\int_M^N A(x)x^{-s-1}\,dx + A(x)x^{-s}\big|_{M^-}^N. \tag{22}$$

In the case that $a_n = 1$ for all n, the function $A(x)$ is $[x]$, the greatest integer function, also called the floor function.

4.2. **Sums and integrals.** Summation by parts is useful for comparing sums and integrals. The sum $\sum_{n=M}^N f(n)$ is expected to be approximately the same as $\int_M^N f(x)\,dx$. Summation by parts gives the difference between them as

$$\int_M^N f(x)d([x] - x) = f(M) + \int_M^N (x - [x])f'(x)\,dx. \tag{23}$$

In the case of $f(x) = x^{-1/2}$ for example the integral is $2N^{1/2} - 1$ and summation by parts gives

$$\begin{aligned}\sum_{n=1}^N n^{-1/2} &= 2N^{1/2} - 1 + \frac{1}{2}\int_1^N ([x] - x)x^{-3/2}\,dx \\ &= 2N^{1/2} + c + \mathrm{O}(N^{1/2})\end{aligned} \tag{24}$$

4.3. **Euler's constant.** Consider $f(x) = 1/x$ and compare the sum and the integral using summation by parts. Euler's constant is defined by the following limit.

As $N \to \infty$

$$\sum_{n=1}^N \frac{1}{n} = \log N + \gamma + \mathrm{O}(1/N) \tag{25}$$

where γ is a constant (called Euler's constant).

Proof. To show the limit exists use summation by parts and the fact that $\log N = \int_1^N dx/x$ to write

$$\sum_{n=1}^N \frac{1}{n} - \log N = 1 + \int_1^N \frac{[x] - x}{x^2}\,dx.$$

Since $0 \le [x] - x < 1$, the integral on the right converges as $N \to \infty$, and we can write

$$\gamma = 1 + \int_1^\infty \frac{[x] - x}{x^2}\,dx.$$

Thus

$$\sum_{n=1}^N \frac{1}{n} - \log N - \gamma = -\int_N^\infty \frac{[x] - x}{x^2}\,dx,$$

and the integral is $\le \displaystyle\int_N^\infty \frac{1}{x^2}\,dx = 1/N$ □

4.4. **Stirling's formula.** Consider $f(x) = \log x$ and compare the sum and the integral using summation by parts. We obtain Stirling's formula.

Stirling's formula. *As* $N \to \infty$

$$\log N! = N \log N - N + \mathrm{O}(\log N).$$

Proof. Write

$$\log N! = \sum_{n=1}^{N} \log n = \int_1^N \log x \, d[x]$$

and

$$N \log N - N + 1 = \int_1^N \log x \, dx.$$

Summation by parts gives

$$\sum_{n=1}^{N} \log n - N \log N + N = 1 + \int_1^N \frac{[x] - x}{x} \, dx.$$

The integral on the right is between 0 and $\log N$. □

4.5. **Hyperbolic sums.** For N a positive integer, the following sets of pairs (m, n) of positive integer are equal.

$$\begin{aligned} &\{mn = k,\, 1 \le k \le N\} \\ &\{1 \le m \le N,\, 1 \le n \le N/m\} \\ &\{1 \le n \le N,\, 1 \le m \le N/n\}\,. \end{aligned} \tag{26}$$

The first expression describes the set as the union of integer points on hyperbolas. The other expressions describe the set as the union of horizontal or vertical lines. It follow that for any function F the following sums are equal.

$$\begin{aligned} &\sum_{1 \le k \le N} \sum_{mn=k} F(m, n) \\ &\sum_{1 \le m \le N} \sum_{1 \le n \le N/m} F(m, n) \\ &\sum_{1 \le n \le N} \sum_{1 \le m \le N/n} F(m, n). \end{aligned} \tag{27}$$

The sum $\sum_{mn=k} F(m, n)$ is of the type found in computing k^{th} coefficient in the product of Dirichlet series, where $F(m, n) = a_m b_n$.

An application for hyperbolic sums is

Asymptotic expression for the divisor function *For a positive integer k let $d(k)$ denote the number of positive divisors of k. Then as $N \to \infty$*

$$\frac{1}{N} \sum_{k=1}^{N} d(k) = \log N + \mathrm{O}(1)$$

More precisely

$$\frac{1}{N} \sum_{k=1}^{N} d(k) = \log N + (2\gamma - 1) + \mathrm{O}(N^{-1/2})$$

where γ is Euler's constant.

The proof uses the fact that the divisor function can be written as a hyperbolic sum,

$$d(k) = \sum_{mn=k} 1.$$

See [7] for the proof.

5. Analytic properties of L functions

In this section we show that the functions $L(s,\chi)$ have analytic continuation to $\operatorname{Re} s > 0$. For a non-trivial character χ the Dirichlet series converges in this region and there are no poles. For the trivial character there is a simple pole at $s = 1$. We show that $L(1,\chi) \neq 0$ which completes the proof of Dirichlet's Theorem.

5.1. **Non trivial characters.** *For χ a non-trivial character the Dirichlet series $L(s,\chi) = \sum_{n=1}^{\infty} \chi(n)n^{-s}$ converges uniformly for $s \geq s_0 > 0$ and $L(s,\chi) - \sum_{n=1}^{N} \chi(n)n^{-s} = \mathrm{O}(N^{-s})$.*

Proof. Use summation by parts. Write $A(x) = \displaystyle\sum_{1 \leq k \leq x} \chi(k)$ and

$$\sum_{n=1}^{N} \chi(n)n^{-s} = \int_1^N x^{-s} dA(x). \tag{28}$$

The function A is bounded. Since a non-trivial character is orthogonal to the trivial character,

$$\sum_{n<k\leq n+q} \chi(k) = 0$$

for any positive integer n. This shows A is a function of period q, and therefore its maximum is on the interval $[0,q]$. For $0 \leq x \leq q$,

$$|A(x)| \leq \sum_{1\leq k\leq x} |\chi(k)| \leq q$$

since $|\chi(k)| \leq 1$ for all k.

Now using summation by parts, (28) can be rewritten as

$$\sum_{n=1}^{N} \chi(n)n^{-s} = A(N)N^{-s} + s\int_1^N x^{-s-1}A(x)\,dx.$$

For $s \geq s_0 > 0$, $|x^{-s-1}A(x)| \leq qx^{-s_0-1}$ and the integral $\displaystyle\int_1^{\infty} x^{-s_0-1}\,dx$ converges for $s_0 > 0$. The first term has absolute value $\leq qN^{-s_0}$ which converges to 0 as $N \to \infty$. Thus $\displaystyle\sum_{n=1}^{N} \chi(n)n^{-s}$ converges uniformly for $s \geq s_0$.

This shows that for $s > 0$,

$$L(s,\chi) = s\int_1^{\infty} x^{-s-1}A(x)\,dx$$

and so

$$
\begin{aligned}
(29) \qquad & L(s,\chi) - \sum_{n=1}^{N} \chi(n) n^{-s} \\
& = -A(N)N^{-s} - s\int_N^\infty A(x)\,dx^{-s} \\
& = \mathrm{O}(N^{-s}).
\end{aligned}
$$

□

The same argument holds for s complex and $\operatorname{Re} s \geq s_0 > 0$. Since the convergence is uniform, the resulting function is holomorphic for $\operatorname{Re} s > 0$.

For χ a non trivial character the Dirichlet series $L(s,\chi) = \sum_{n=1}^\infty \chi(n)n^{-s}$ converges to a holomorphic function for $\operatorname{Re} s > 0$.

We now know that $\lim_{s\to 1} L(s,\chi) = L(1,\chi)$ for a non-trivial character.

5.2. **The trivial character.** The function $L(s,\chi_0)$ for the trivial character is an entire function times the Riemann zeta function ζ. To see this, look at the Euler products for $L(s,\chi_0)$ and $\zeta(s)$.

$$
L(s,\chi_0) = \prod_p (1-\chi_0(p)p^{-s})^{-1} \qquad \zeta(s) = \prod_p (1-p^{-s})^{-1}.
$$

Since for a prime p, $\chi_0(p) = 1$ unless $p|q$, in which case $\chi_0(p) = 0$, it follows that

$$
L(s,\chi_0) = \zeta(s)\prod_{p|q}(1-p^{-s})^{-1}. \tag{30}
$$

Note that the product on the right is a finite product and therefore and entire function. It is also not zero at $s = 1$.

We show that the function ζ can be analytically continued to a function which has a simple pole at $s = 1$.

The function $\zeta(s) - 1/(s-1)$ can be analytically continued to a function holomorphic in $\operatorname{Re} s > 0$. If follow that ζ can be extended to a meromorphic function there with simple pole at $s = 1$.

Proof. Using summation by parts.

$$
\sum_{n=1}^N n^{-s} - \int_1^N x^{-s}\,dx = 1 + s\int_1^N x^{-s-1}([x]-x)\,dx.
$$

Since the integrand on the right is $\leq x^{-\operatorname{Re} s - 1}$ in absolute value, the integral converges as $N \to \infty$ uniformly in $\operatorname{Re} s \geq s_0 > 0$ to a holomorphic function. Letting $N \to \infty$, for $\operatorname{Re} s > 1$ get

$$
\zeta(s) = \frac{1}{s-1} + 1 + s\int_1^\infty x^{-s-1}([x]-x)\,dx \tag{31}
$$

and the expression on the right is meromorphic for $\operatorname{Re} s > 0$. □

So now by (30), $L(s,\chi_0)$ has a simple pole at $s = 1$.

5.3. **Non vanishing of L function at $s = 1$.** We show that $L(1,\chi) \neq 0$ if χ is a non-trivial character.

If χ has values which are not real it is called a *complex character*. In this case the proof is easy

For χ a complex Dirichlet character, $L(1,\chi) \neq 0$.

Proof. Form the product of all functions $L(s,\chi)$ for Dirichlet characters χ mod q,

$$L(s) = \prod_{\chi} L(s,\chi). \tag{32}$$

Since

$$\log L(s,\chi) = -\sum_{p} \log(1 - \chi(p)p^{-s}) = \sum_{p} \sum_{k=1}^{\infty} \frac{1}{k}\chi(p^k)p^{-sk}$$

then

$$\begin{aligned} \log L(s) &= \sum_{\chi} \sum_{p} \sum_{k=1}^{\infty} \frac{1}{k}\chi(p^k)p^{-sk} \\ &= \sum_{p} \sum_{k=1}^{\infty} \sum_{\chi} \frac{1}{k}\chi(p^k)p^{-sk} \end{aligned} \tag{33}$$

By the orthogonality of characters

$$\sum_{\chi} \overline{\chi(\ell)}\chi(m) = \delta_\ell(m),$$

where $\delta_\ell(m)$ is defined by (17). Letting $\ell = 1$, and $m = p^k$,

$$\sum_{\chi} \chi(p^k) = \delta_1(p^k).$$

Substituting in (33),

$$\log L(s) = \sum_{p} \sum_{k=1}^{\infty} \frac{1}{k}\delta_1(p^k)p^{-sk} \tag{34}$$

showing that for s real $\log L(s) \geq 0$ so $L(s) \geq 1$.

If χ is a complex character, then so is $\overline{\chi}$. If $L(1,\chi) = 0$ then also $L(1,\overline{\chi}) = 0$, and in the product (32) one of the factors has a simple pole and two have zeros at 1. This implies that $L(1) = 0$. But this contradicts the fact that $L(s) \geq 1$ for s real. □

If all values of χ are real, it is said to be a *real character*.

For χ a non-trivial real Dirichlet character, $L(1,\chi) \neq 0$.

Proof. The proof uses the follwing sum along hyperbolas

$$S_N = \sum_{k=1}^{N} \sum_{mn=k} \frac{\chi(n)}{(nm)^{1/2}}. \tag{35}$$

That $L(1,\chi) \neq 0$ follows from the two statements

$$S_N \geq c \log N \text{ for some constant } c > 0 \tag{36}$$

$$S_N = 2N^{1/2}L(1,\chi) + \mathrm{O}(1) \text{ as } N \to \infty \tag{37}$$

since if $L(1, \chi) = 0$, (37) says that S_N is bounded as $N \to \infty$ which contradicts (36).

Proof of (36) Write

$$S_N = \sum_{k=1}^{N} \sum_{n|k} \chi(n)$$

Now using

$$\sum_{n|k} \chi(n) \geq \begin{cases} 0 & \text{for all } k \\ 1 & \text{if } k = \ell^2 \text{ for some } \ell \in \mathbb{Z}. \end{cases} \tag{38}$$

(this is shown below) get

$$S_N \geq \sum_{\ell \leq N^{1/2}} \frac{1}{\ell} = \log(N^{1/2}) + \mathrm{O}(1)$$

where the right inequality is from (25). This proves (36).

To prove (38) look at the prime factorization $k = p_1^{a_1} \cdots p_n^{a_n}$ of k. The divisors of k are $p_1^{b_1} \cdots p_n^{b_n}$ for $0 \leq b_j \leq a_j$. By the multiplicative property of characters we can factor $\sum_{n|k} \chi(n)$ as

$$\sum_{n|k} \chi(n) = \prod_{j=1}^{n} \left(\chi(1) + \chi(p_j) + \chi(p_j^2) + \cdots \chi(p_n^{a_n}) \right). \tag{39}$$

But for any p and a,

$$\chi(1) + \chi(p) + \chi(p^2) + \cdots \chi(p^a) = \begin{cases} a+1 & \text{if } \chi(p) = 1 \\ 1 & \text{if } \chi(p) = -1 \text{ and } a \text{ is even} \\ 0 & \text{if } \chi(p) = -1 \text{ and } a \text{ is odd} \\ 1 & \text{if } \chi(p) = 0, \text{ that is } p|q \end{cases}$$

Now if is k is a square, all of the a_j in (39) are even and each factor is ≥ 1 . In any case each factor is ≥ 0. This proves (38).

Proof of (37) First prove the following two inequalities for integers $a < b$, as $a \to \infty$

$$\sum_{n=a}^{b} \frac{\chi(n)}{n^{1/2}} = \mathrm{O}(a^{-1/2}) \tag{40}$$

$$\sum_{n=a}^{b} \frac{\chi(n)}{n} = \mathrm{O}(a^{-1}) \tag{41}$$

Equation (40) follow from (29) since the expression can be written as

$$\begin{aligned} L(\chi, 1/2) - \sum_{n=1}^{a} \chi(n) n^{-1/2} - \left(L(\chi, 1/2) - \sum_{n=1}^{b} \chi(n) n^{-1/2} \right) \\ = \mathrm{O}(a^{-1/2}) + \mathrm{O}(b^{-1/2}). \end{aligned}$$

Now the result follows since for $b > a$, $b^{-1/2} < a^{-1/2}$.

The proof of (41) also follows from (29) by a similar argument.

For the rest of the proof see [7]. $\square$

6. Prime counting functions

There are an infinite number of primes and an infinite number of primes of the form $nq + \ell$ for q and ℓ relatively prime. What about a more precise count of primes? It was first observed by Gauss and Legendre that the number of primes less than or equal to a positive number x is approximately $x/\log x$. Let

$$\pi(x) = \sum_{p \le x} 1 = \text{ the number of primes} \le x \tag{42}$$

be the prime counting function. The Prime Number Theorem (PNT) conjectured by Gauss and Legendre is

$$\pi(x) \sim \frac{x}{\log x} \text{ as } x \to \infty, \tag{43}$$

which is short a way of saying that

$$\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 1. \tag{44}$$

To prove PNT we must study more carefully the behavior of $\zeta(s)$ at $s = 1$.

6.1. **Generating functions and counting primes.** In keeping with the spirit of using generating functions, consider a Dirichlet series $\sum_n a_n n^{-s}$ with coefficients $a_n = 1$ for n a prime and 0 otherwise. This is the function $\sum_p p^{-s}$ that was studied to give the analytic proof that there are an infinite number of primes. The prime counting function (42) is the partial sum of the coefficients and so using summation by parts,

$$\begin{aligned} \sum_p p^{-s} &= \int_0^\infty x^{-s}\, d\pi(x) \\ &= s\int_0^\infty x^{-s-1}\pi(x)\, dx. \end{aligned} \tag{45}$$

Some other prime counting functions are more convenient for proving PNT (43). Consider the derivative $\sum_p (\log p)p^{-s}$ of (45). The partial sum of the coefficients is given by the function

$$\vartheta(x) = \sum_{p\le x} \log p. \tag{46}$$

The function ϑ gives another way to count primes, weighted by their log. Summation by parts gives

$$\begin{aligned} \sum_p (\log p)p^{-s} &= \int_0^\infty x^{-s}\, d\vartheta(x) \\ &= s\int_0^\infty x^{-s-1}\vartheta(x)\, dx \end{aligned} \tag{47}$$

since clearly $\vartheta(x) \le x\log x$, so $x^{-s}\vartheta(x) \to 0$ as $x \to \infty$ for $\operatorname{Re} s > 1$.

The Prime Number Theorem is equivalent to

$$\vartheta(x) \sim x \text{ as } x \to \infty. \tag{48}$$

Proof. This follows from the fact that for every $\epsilon > 0$,

$$(1-\epsilon)\pi(x)\log x + \mathrm{O}(x^{1-\epsilon})\log x \le \vartheta(x) \le \pi(x)\log x. \tag{49}$$

The inequality on the right is easy. To get the one on the left, first note that it is clear that $\pi(x) \le x$. Then for $\epsilon > 0$,

$$\begin{aligned}\vartheta(x) \ge \sum_{x^{1-\epsilon}<p\le x} \log p &\ge \log x^{1-\epsilon} \sum_{x^{1-\epsilon}<p\le x} 1 \\ &= \log x^{1-\epsilon}\left(\pi(x) - \pi(x^{1-\epsilon})\right) \\ &= \pi(x)(1-\epsilon)\log x + \mathrm{O}(x^{1-\epsilon})\log x.\end{aligned}$$

So if $\lim_{x\to\infty}\vartheta(x)/x$ exists, (49) and the fact that $\lim_{x\to\infty} x^{-\epsilon}\log x = 0$ shows that

$$(1-\epsilon)\varliminf_{x\to\infty}\frac{\pi(x)\log x}{x} \le \lim_{x\to\infty}\frac{\vartheta(x)}{x} \le \varlimsup_{x\to\infty}\frac{\pi(x)\log x}{x}$$

for every $\epsilon > 0$. So $\lim_{x\to\infty}(\pi(x)\log x)/x$ exists and

$$\lim_{x\to\infty}\frac{\pi(x)\log x}{x} = \lim_{x\to\infty}\frac{\vartheta(x)}{x}.$$

A similar argument proves the above equation assuming the limit on the left exists. □

Another prime counting function is obtained by looking at the Dirichlet series for ζ'/ζ. By the Euler factorization of ζ,

$$\begin{aligned}-\frac{\zeta'(s)}{\zeta(s)} &= \sum_p \frac{p^{-s}\log p}{1-p^{-s}} \\ &= \sum_p \log p \sum_{k=1}^{\infty} p^{-ks} \\ &= \sum_{n=1}^{\infty}\Lambda(n)n^{-s}\end{aligned} \tag{50}$$

for $\operatorname{Re} s > 1$, where Λ is the *Von Mangoldt function* defined by

$$\Lambda(n) = \begin{cases}\log p & \text{if } n = p^k \\ 0 & \text{otherwise.}\end{cases} \tag{51}$$

Define the partial sums of the coefficients of (50) by

$$\psi(x) = \sum_{n\le x}\Lambda(k). \tag{52}$$

Using this new prime counting function we have a new statement of the Prime Number Theorem.

The Prime Number Theorem is equivalent to

$$\psi(x) \sim x \text{ as } x \to \infty, \tag{53}$$

Proof.

$$\psi(x) - \vartheta(x) = \sum_{\substack{p^k \le x \\ 2 \le k}} \log p$$

$$\le \sum_{p \le \sqrt{x}} \log p \le \sqrt{x} \log x$$

Now $(\sqrt{x} \log x)/x$ goes to 0 as $x \to \infty$, so by (48), (53) is another statement of PNT. □

The functions $\psi(x)$ and $\vartheta(x)$ are generated by partial sums of the coefficients of $-\zeta'(s)/\zeta(s)$ and $\sum_p (\log p) p^{-s}$ respectively. These functions are closely related.

The function

$$\sum_p (\log p) p^{-s} + \frac{\zeta'(s)}{\zeta(s)} \tag{54}$$

has analytic continuation to a function holomorphic in $\mathrm{Re}\, s > 1/2$.

Proof. The Euler factorization of ζ gives

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{\log p}{p^s - 1}.$$

so (54) is given by

$$-\sum_p \frac{\log p}{p^s(p^s - 1)}.$$

Comparing this series to $\sum_p p^{-2\mathrm{Re}\, s + \epsilon}$, it converges for $\mathrm{Re}\, s > (1+\epsilon)/2$. Since this is true for every $\epsilon > 0$ this completes the proof. □

This shows that both functions $-\zeta'(s)/\zeta(s)$ and $\sum_p (\log p) p^{-s}$ have the same singular part $1/(s-1)$ at $s = 1$.

6.2. **Outline of proof of PNT.** We can now give an argument that suggests why PNT is true, but which does not give a proof. The previous paragraph shows that the function defined in $\mathrm{Re}\, s > 1$ by

$$\sum_p (\log p) p^{-s} - \frac{1}{s-1} \tag{55}$$

has analytic continuation to a function holomorphic in a neighborhood of $s = 1$. Use summation by parts (47) and the fact that $\int_1^\infty x^{-s}\, dx = 1/(s-1)$ to write (55) for $\mathrm{Re}\, s > 1$ as

$$\int_1^\infty \left(\frac{s\vartheta(x)}{x} - 1 \right) x^{-s}\, dx. \tag{56}$$

If we could take the limit under the integral sign as $s \to 1^+$ we would get that the integral

$$\int_1^\infty \left(\frac{\vartheta(x)}{x} - 1 \right) \frac{dx}{x} \tag{57}$$

converges. This integral is of the form

$$\int_1^\infty f(x) x^{-1}\, dx = \int_0^\infty f(e^t)\, dt$$

where $f(x) = \vartheta(x)/x - 1$. Since the integral converges, probably $f(e^t) \to 0$ as $t \to \infty$ and this is PNT. It is not true in general that if $\int_0^\infty g(t)\,dt$ converges then $g(t) \to 0$ as $t \to \infty$. There are easy counterexamples. In this case, however, the function g is regular enough that it is true. The fact that $\vartheta(x)$ is increasing is the key.

If

$$\int_1^\infty \left(\frac{\vartheta(x)}{x} - 1\right)\frac{dx}{x} \tag{58}$$

converges then

$$\lim_{x\to\infty}\left(\frac{\vartheta(x)}{x} - 1\right) = 0 \tag{59}$$

and this proves PNT.

Proof. This proof is taken from [8]. If (59) does not hold then for some $\epsilon > 0$ there is a sequence x_n with limit ∞ such that for every n

$$\frac{\vartheta(x_n)}{x_n} > 1 + \epsilon \quad \text{or} \quad \frac{\vartheta(x_n)}{x_n} < 1 - \epsilon.$$

One of these inequalities must hold for infinitely many n.

Now let $\lambda = 1 + \epsilon$. Suppose for some $\lambda > 1$ there are arbitrarily large x with $\vartheta(x) \geq \lambda x$. Since ϑ is non-decreasing,

$$\begin{aligned}\int_x^{\lambda x} \frac{\vartheta(t) - t}{t^2}\,dt &\geq \int_x^{\lambda x} \frac{\lambda x - t}{t^2}\,dt \\ &= \int_1^{\lambda} \frac{\lambda - t}{t^2}\,dt > 0.\end{aligned}$$

Note the last expression is independent of x. This contradicts the convergence of the integral (58) since if the integral converges,

$$\lim_{x\to\infty} \int_x^{\lambda x} \frac{\vartheta(t) - t}{t^2}\,dt = 0.$$

Likewise let $\lambda = 1 - \epsilon$. Suppose for some $\lambda < 1$ there are arbitrarily large x with $\vartheta(x) \leq \lambda x$. Since ϑ is non-decreasing,

$$\begin{aligned}\int_{\lambda x}^{x} \frac{\vartheta(t) - t}{t^2}\,dt &\leq \int_{\lambda x}^{x} \frac{\lambda x - t}{t^2}\,dt \\ &= \int_{\lambda}^{1} \frac{\lambda - t}{t^2}\,dt < 0\end{aligned}$$

which again contradicts the convergence of the integral (58). $\square$

These same arguments can be made with $\vartheta(x)$ replaced by $\psi(x)$.

So PNT reduces to proving that the integral $\displaystyle\int_1^\infty \left(\frac{\vartheta(x)}{x} - 1\right)\frac{dx}{x}$ converges. This is the hard part and will require Newman's Analytic Theorem proved in section 8

6.3. **The Möbius function.** Another prime counting function is obtained by looking at partial sums of the coefficients of $1/\zeta$. By the Euler factorization,

$$\frac{1}{\zeta(s)} = \prod_p (1 - p^{-s}) = \sum_{n=1}^{\infty} \mu(n) n^{-s}, \tag{60}$$

where the coefficients are given by *the Möbius function*

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^k & \text{if } n = p_1 \cdots p_k \text{ for distinct primes,} \\ 0 & \text{otherwise.} \end{cases} \tag{61}$$

The Möbius function is interesting because it has values $-1, 0$ or 1 and appears to be quite random in a way explained later. The partial sums of the coefficients of the series for $1/\zeta$ are denoted by

$$M(x) = \sum_{n \le x} \mu(n). \tag{62}$$

We show later that PNT is equivalent to $M(x)/x \to 0$ as $x \to \infty$. So it is known to be random in the sense that the average goes to zero.

Relationship between the Möbius function and von Mangoldt function and the sequence $\log n$ can be seen by multiplying the various generating functions. Recall the following Dirichlet series.

$$\begin{aligned} -\frac{\zeta'(s)}{\zeta(s)} &= \sum_{n=1}^{\infty} \Lambda(n) n^{-s} \\ \frac{1}{\zeta(s)} &= \sum_{n=1}^{\infty} \mu(n) n^{-s} \\ \zeta(s) &= \sum_{n=1}^{\infty} n^{-s} \\ -\zeta'(s) &= \sum_{n=1}^{\infty} (\log n) n^{-s} \end{aligned} \tag{63}$$

Since the fourth equation is the first times the third, get by equating coefficients of the Dirichlet series

$$\log n = \sum_{d|n} \Lambda(d). \tag{64}$$

Since the first equation is the second times the third, get by equating coefficients of the Dirichlet series

$$\Lambda(k) = \sum_{mn=k} \mu(m) \log n. \tag{65}$$

Equations (64) and (64) can be considered as inverses of each other.

In proving equation (64) we have assumed that the coefficients of Dirichlet series are unique, but this equation, like the Euler factorization, is based simply on the the fact that every integer can be uniquely factored into primes. To prove (64) using this fact, write $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. Then the only d for $\Lambda(d)$ is not zero are $d = p_j^{\beta}$ for $1 \le \beta \le \alpha_j$ and (64) follows directly. Equation (65) can be proved similarly.

TABLE 2. Dirichlet series related to $\zeta(s)$ and partial sums of coefficients. Coefficients not specified are 0. Primes $p_1, \ldots, p_k$ are distinct. Coefficients can be easily found from the Euler factorization $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$

series $\sum a_n n^{-s}$	coeff. a_n	coeff. name	part. sum $\sum_{n \le x} a_n$	part. sum name
$\zeta(s)$	1		$[x]$	floor
$1/\zeta(s)$	$(-1)^k$ if $n = p_1 \cdots p_k$	$\mu(n)$ Möbius	$M(x)$	Mertens
$-\zeta'(s)$	$\log n$		$\log([x]!)$	$T(x)$
$-\zeta'(s)/\zeta(s)$	$\log p$ if $n = p^k$	$\Lambda(n)$ von Mangoldt	$\psi(x)$	Chebychev
$\sum_p (\log p) p^{-s}$	$\log p$ if $n = p$		$\vartheta(x)$	
$\log \zeta(s)$	$\Lambda(n)/\log n$		$J(x)$	
$\sum_p p^{-s}$	1 if $n = p$		$\pi(x)$	prime counting
$\zeta(2s)/\zeta(s)$	$(-1)^{a_1 + \cdots + a_k}$ if $n = p_1^{a_1} \cdots p_k^{a_k}$	$\lambda(n)$ Liouville	$M_1(x)$	

7. Chebyshev's estimates

Chebyshev studied the function $\psi(x) = \sum_{n \le x} \Lambda(n)$. The Prime Number Theorem is equivalent to $\psi(x)/x \to 1$ as $x \to \infty$. Chebyshev was able to show without complex analysis that for sufficiently large x

$$A < \psi(x)/x < 6A/5 \tag{66}$$

where $A = .921\ldots$, and $6A/5 = 1.106\ldots$. The technique is interesting because it uses only elementary analysis and Stirling's formula. We will give the proof in this section. This material is taken from [4].

7.1. **An easy upper estimate.** As a warm up, we prove something easier. The proof does not use any analysis, just some algebra. We show that the order of magnitude of $\vartheta(x)$ is less that x in the sense that there is a constant C such that $\vartheta(x)/x \le C$ for large x, that is, $\vartheta(x) = \mathrm{O}(x)$. This will be an important hypothesis used in the proof of Newman's Analytic Theorem, the key step in the proof of PNT.

Theorem.

$$\frac{\vartheta(x)}{x} \le 4 \log 2 \tag{67}$$

for x sufficiently large.

Proof. For any integer $n \geq 1$ see by the binomial expansion of $(1+1)^{2n}$ that

$$\begin{aligned} 2^{2n} &= \sum_{k=0}^{2n} \binom{2n}{k} > \binom{2n}{n} \\ &= \frac{(2n)(2n-1)\dots(n+1)}{n!} \\ &\geq \prod_{n<p\leq 2n} p. \end{aligned} \tag{68}$$

The latter inequality holds since any prime p, $n < p \leq 2n$ is one of the factors of the numerator but does not divide any factor of the denominator, and therefore must divide $\binom{2n}{n}$ which is an integer.

Taking the log of (68) gives

$$(2\log 2)n \geq \sum_{n<p\leq 2n} \log p = \vartheta(2n) - \vartheta(n).$$

Letting $n = 2^k$ get

$$(2\log 2)2^k \geq \vartheta(2^{k+1}) - \vartheta(2^k).$$

It follows that

$$\begin{aligned} \vartheta(2^{n+1}) &= \sum_{k=0}^{n} [\vartheta(2^{k+1}) - \vartheta(2^k)] \\ &\leq 2\log 2 \sum_{k=0}^{n} 2^k = (2\log 2)(2^{n+1} - 1) \\ &\leq (2\log 2)2^{n+1}. \end{aligned}$$

Now if $x > 1$ then for some k, $2^k \leq x \leq 2^{k+1}$ and

$$\vartheta(x) \leq \vartheta(2^{k+1}) \leq (2\log 2)2^{k+1} \leq (4\log 2)x.$$

Thus if $C = 4\log 2$ the inequality $\vartheta(x) \leq Cx$ holds. With a little more work it can be shown using this method that any $C > 2\log 2$ works.

$\square$

7.2. **Upper and lower estimates.** Now proceed to the proof of (66). Chebychev used the formulas (64) and (65).

Let

$$T(x) = \log([x]!) = \sum_{k\leq x} \log k.$$

Summing both sides of (64) over $k \leq x$, and reordering the sum over hyperbolas to the sum over vertical lines, get

$$\begin{aligned} T(x) &= \sum_{k\leq x} \log k = \sum_{k\leq x} \sum_{mn=k} \Lambda(n) \\ &= \sum_{m\leq x} \sum_{n\leq x/m} \Lambda(n) \\ &= \sum_{m\leq x} \psi(x/m). \end{aligned} \tag{69}$$

Summing both sides of (65) over $k \le x$ and reordering the sum over hyperbolas to the sum over vertical lines get

$$\begin{aligned} \psi(x) &= \sum_{k\le x} \Lambda(k) = \sum_{k\le x} \sum_{mn=k} \mu(m)\log n \\ &= \sum_{m\le x} \mu(m) \sum_{n\le x/m} \log n \\ &= \sum_{m\le x} \mu(m) T(x/m). \end{aligned} \tag{70}$$

The absolute values, $T(x/m)$, of the terms in the last sum can be estimated using Stirling's formula $T(x) = x\log x - x + \mathrm{O}(\log x)$ to try to get an estimate of $\psi(x)$ as $x \to \infty$. The problem is that the coefficients $\mu(m)$ are a seemingly random sequence chosen from $-1, 0, 1$, and therefore many cancellations occur.

The method of Chebyshev was to consider instead of the right hand side of (70) the sum

$$\begin{aligned} &T(x) - T(x/2) - T(x/3) \\ &- T(x/5) + T(x/30) = \sum_{m\in S} \nu_m T(x/m) \end{aligned} \tag{71}$$

where S is the set of integers $\{1, 2, 3, 5, 30\}$ and where $\nu_1 = 1$, $\nu_2 = -1$, $\nu_3 = -1$, $\nu_5 = -1$ and $\nu_{30} = 1$. Substituting (69) in the sum (71) get

$$\begin{aligned} \sum_{m\in S} \nu_m T(x/m) &= \sum_{m\in S} \nu_m \sum_{n\le x/m} \psi\left(x/(mn)\right) \\ &= \sum_{k\le x} \sum_{\substack{m\in S \\ m|k}} \nu_m \psi(x/k) \\ &= \sum_{k\le x} A_k \psi(x/k) \end{aligned} \tag{72}$$

where

$$A_k = \sum_{\substack{m\in S \\ m|k}} \nu_m. \tag{73}$$

For $m \in S$ $m|30$, so $m|k$ if and only if $m|(k+30)$ and it follows that $A_k = A_{k+30}$. Thus A_k is completely determined by its values for $k = 1, 2, \ldots 30$ given in table 3.

Table 3. Table of A_k given by (73)

k	1	2	3	4	5	6	7	8	9	10
A_k	1	0	0	0	0	-1	1	0	0	-1
k	11	12	13	14	15	16	17	18	19	20
A_k	1	-1	1	0	-1	0	1	-1	1	-1
k	21	22	23	24	25	26	27	28	29	30
A_k	0	0	1	-1	0	0	0	0	1	-1

From table 3 and the fact that ψ is an increasing function, we see that the series $\sum_{k=1}^{\infty} A_k \psi(x/k)$ is an alternating series of the form $\sum_{j=1}^{\infty} (-1)^{j+1} a_j$ with

a_j a decreasing sequence of non-negative numbers. In the latter case the sum σ satisfies $a_1 - a_2 < \sigma < a_1$. Thus

$$\psi(x) - \psi(x/6) \leq \sum_{m\in S} \nu_m T(x/m) \leq \psi(x). \tag{74}$$

Using Stirling's formula $T(x) = x\log x - x + \mathrm{O}(\log x)$ and the fact that $\sum_{n\in S} \nu_n/n = 0$, get

$$\sum_{m\in S} \nu_m T(x/m) = Ax + \mathrm{O}(\log x) \tag{75}$$

where

$$A = \sum_{m\in S} \nu_m \frac{\log m}{m} = .921\ldots$$

Now (74) becomes

$$\psi(x) - \psi(x/6) \leq Ax + \mathrm{O}(\log x) \leq \psi(x). \tag{76}$$

and this gives the left hand inequality in (66).

To get the right hand inequality of (66), substitute $\frac{x}{6^k}$ in the left inequality of (76), and sum getting

$$\psi\left(x\right) = \sum_{k\leq \frac{\log x}{\log 6}} \left(\psi\left(\frac{x}{6^k}\right) - \psi\left(\frac{x}{6^{k+1}}\right)\right) \tag{77}$$

since if $k > \frac{\log x}{\log 6}$ then $\frac{x}{6^k} < 1$ and $\psi\left(\frac{x}{6^k}\right) = 0$. Hence by (76)

$$\begin{aligned}\psi\left(x\right) &\leq Ax \sum_{k\leq \frac{\log x}{\log 6}} 6^{-k} + \mathrm{O}(\log x)\frac{\log x}{\log 6}\\ &\leq Ax\sum_{k=0}^{\infty} 6^{-k} + \mathrm{O}((\log x)^2)\\ &= \frac{5}{6}Ax + \mathrm{O}((\log x)^2)\end{aligned} \tag{78}$$

and the right hand inequality of (66) follows.

8. Proof of the Prime Number Theorem

The easiest proof of PNT follows from a study of $\zeta(s)$ using complex variables theory. There are "elementary" proofs which do not use complex variables, but they are more complicated. We will show that $\zeta(s)$ has no zeros on $\mathrm{Re}\, s = 1$ and that this fact is equivalent to PNT. This explains why Chebychev could not complete the proof of PNT, since he was looking only along the real axis.

8.1. **PNT and zeros of ζ.** Recall from (54) that the function

$$\Phi(s) = \sum_p \frac{\log p}{p^s} \tag{79}$$

can be continued to $\mathrm{Re}\, s > 1/2$ and that $\Phi(s) + \dfrac{\zeta'(s)}{\zeta(s)}$ is holomorphic in $\mathrm{Re}\, s > 1/2$. The function $\frac{\zeta'(s)}{\zeta(s)}$ has simple poles with residues equal to the order of the zero of $\zeta(s)$ at the poles. Thus if s_0 is a zero of ζ of order h then $\lim\limits_{s\to s_0}(s - s_0)\Phi(s) = -h$.

First we show

The Prime Number Theorem, $\vartheta(x) \sim x$ as $x \to \infty$, implies there are no zeros of $\zeta(s)$ on $\operatorname{Re} s = 1$.

Proof. In sections 6.1 and 6.2 it was shown that for $\operatorname{Re} s > 1$,

$$\begin{aligned} \int_1^\infty (\vartheta(x) - x)x^{-s-1}\,dx &= \frac{\Phi(s)}{s} - \frac{1}{s-1} \\ &= -\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1} + h(s) \end{aligned} \tag{80}$$

where $h(s)$ is holomorphic on $\operatorname{Re} s > 1/2$. Denote this meromorphic function in $\operatorname{Re} s > 1/2$ by $f(s)$. We show $f(s)$ has no poles on $\operatorname{Re} s = 1$ and therefore $\zeta(s)$ has no zeros there. For $\epsilon > 0$ choose M such that

$$|\vartheta(x) - x| \le \epsilon x$$

for $x > M$. Then if $s = \sigma + i\tau$, $\sigma > 1$,

$$\begin{aligned} |f(s)| &\le \int_1^M |\vartheta(x) - x|x^{-\sigma-1}\,dx + \epsilon\int_M^\infty x^{-\sigma}\,dx \\ &= g(\sigma) + \epsilon\frac{M^{1-\sigma}}{\sigma-1} \end{aligned} \tag{81}$$

where $g(\sigma)$ is continuous at $\sigma = 1$. Let $s_0 = 1 + i\tau$, then

$$\left|\lim_{\substack{s\to s_0\\ \operatorname{Im} s=\tau}} (s - s_0)f(s)\right| = \lim_{\sigma\to 1}(\sigma - 1)|f(\sigma + i\tau)| \le \epsilon.$$

Since ϵ is arbitrary,

$$\lim_{\substack{s\to s_0\\ \operatorname{Im} s=\tau}} (s - s_0)f(s) = 0. \tag{82}$$

So $\zeta'(s)/\zeta(s)+1/(s-1)$ has no singularity at s_0, since at a singularity the limit (82) is ∞ or non-zero. Since τ is arbitrary, $\zeta'(s)/\zeta(s)$ has no singularities on $\operatorname{Re} s = 1$ except at $s = 1$, so $\zeta(s)$ has no zeros on $\operatorname{Re} s = 1$. □

8.2. **There are no zeros of $\zeta(s)$ on $\operatorname{Re} s = 1$.** In this section we prove that

There are no zeros of $\zeta(s)$ on $\operatorname{Re} s = 1$

Proof. For any positive real number p and any real number α, $p^{i\alpha} + p^{-i\alpha}$ is real so

$$\begin{aligned} 0 \le (p^{i\alpha/2} + p^{-i\alpha/2})^4 &= \sum_{k=0}^4 \binom{4}{k} p^{ik\alpha/2}p^{-i(4-k)\alpha/2} \\ &= p^{-2i\alpha} + 4p^{-i\alpha} + 6 + 4p^{i\alpha} + p^{2i\alpha}. \end{aligned} \tag{83}$$

Now let

$$\Phi(s) = \sum_p \frac{\log p}{p^s}. \tag{84}$$

For σ real, multiply equation (83) by $p^{-\sigma}\log p$ and sum over all primes p to get

$$\begin{aligned} 0 \le{}& \Phi(\sigma + 2i\alpha) + 4\Phi(\sigma + i\alpha) + 6\Phi(\sigma) \\ &+ 4\Phi(\sigma - i\alpha) + \Phi(\sigma - 2i\alpha). \end{aligned} \tag{85}$$

Suppose now that for $\alpha \neq 0$, ζ has a zero of order μ at $1+i\alpha$ and order ν at $1+i2\alpha$. Since the coefficients of ζ are real, ζ also has a zero of order μ at $1-i\alpha$ and order ν at $1-i2\alpha$. Then Φ has simple pole with residue μ at $1 \pm i\alpha$ and residue ν at $1 \pm i2\alpha$. Also $\Phi(s)$ has a simple pole with residue 1 at $s=0$. Multiplying equation (85) by $\sigma - 1$ and letting $\sigma \to 1$ get $0 \leq -2\nu - 8\mu + 6$. Since μ is a positive integer, conclude that $\mu = 0$. Thus there is no zero at $1+i\alpha$. Since $\alpha \neq 0$ is arbitrary, there is no zero of ζ on $\operatorname{Re} s = 1$.

□

8.3. **Newman's Analytic Theorem.** Using Newman's Analytic Theorem we prove that no zeros of ζ on $\operatorname{Re} s = 1$ implies PNT. Newman's Theorem belongs to a class of theorems called *Tauberian theorems* which deal with taking the limit under the sum or integral sign. Newman's Tauberian theorem has weaker hypotheses that the Wiener-Ikehara Tauberian theorem often used to prove PNT, but using Chebyshev's estimate $\vartheta(x) = \mathrm{O}(x)$, Newman's Theorem proves PNT.

Newman's Analytic Theorem I: Let $f(t)$ be a bounded and locally integrable funtion on $[0.\infty)$ and suppose that the function $g(z) = \int_0^\infty f(t)e^{-zt}\,dt$, $\operatorname{Re} z > 0$ extends holomorphically to $\operatorname{Re} z \geq 0$. Then $\int_0^\infty f(t)\,dt$ exists (and equals $g(0)$).

Newman's Analytic Theorem essentially says that under the conditions of the theorem, the limit can be taken under the integral sign,

$$
\begin{aligned}
g(0) &= \lim_{z\to 0^+} \int_0^\infty f(t)e^{-zt}\,dt \\
&= \int_0^\infty \lim_{z\to 0^+} f(t)e^{-zt}\,dt \\
&= \int_0^\infty f(t)\,dt.
\end{aligned}
\tag{86}
$$

implying that the last integral exists.

Using the change of variables $x = e^t$ and $s = z+1$, Newman's Theorem becomes:

Newman's Analytic Theorem II: Let $f(x)$ be a bounded and locally integrable funtion on $[1,\infty)$ and suppose that the function $g(s) = \int_0^\infty f(x)x^{-s}\,dx$, $\operatorname{Re} s > 1$ extends holomorphically to $\operatorname{Re} s \geq 1$. Then $\int_1^\infty f(x)x^{-1}\,dx$ exists (and equals $g(1)$).

To prove PNT, use the second form of the Analytic Theorem with $f(x) = \dfrac{\vartheta(x)}{x} - 1$. The Analytic Theorem is a little easier to prove in the first form.

Note: The integral $g(z) = \int_0^\infty f(t)e^{-zt}\,dt$ is called the *Fourier transform* of f, and the integral $g(s) = \int_0^\infty f(x)x^{-s}\,dx$ is called the *Mellin transform* of f.

Proof. We prove of Newman's Analytic Theorem I. The proof uses the Cauchy integral theorem in an extremely clever way.

Let

$$g_T(z) = \int_0^T f(t)e^{-zt}\,dt$$

We must show that

$$\lim_{T\to\infty} g_T(0) = g(0). \tag{87}$$

The difference $g(0) - g_T(0)$ is estimated using Cauchy's theorem. Fix $R > 0$ and $\delta > 0$ such that $g(z)$ is holomorphic in the region

$$\mathcal{R} = \{z \mid |z| \leq R \text{ and } \operatorname{Re} z \geq -\delta\}$$

and let C be the boundary of $\mathcal{R}$. Since 0 is in the interior, by Cauchy's Theorem,

$$g(0) - g_T(0) = \frac{1}{2\pi i} \int_C (g(z) - g_T(z))\, e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z}. \tag{88}$$

Break the integral into three parts. Let

$$C_+ = C \cap \{z \mid \operatorname{Re} z > 0\}$$

and

$$C_- = C \cap \{z \mid \operatorname{Re} z < 0\}$$

so that $C = C_+ + C_-$, and C_+ is a semicircle. Then

$$g(0) - g_T(0) = I_1 + I_2 + I_3 \tag{89}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{2\pi i} \int_{C_+} (g(z) - g_T(z))\, e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z} \\
I_2 &= -\frac{1}{2\pi i} \int_{C_-} g_T(z) e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z} \\
I_3 &= \frac{1}{2\pi i} \int_{C_-} g(z) e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z}.
\end{aligned}$$

We estimate the integrals I_1, I_2, I_3 separately.

To estimate I_1, suppose $|f(t)| \leq B$ for all t and observe that on C_+ , $|z| = R$ and

$$\begin{aligned}
|g(z) - g_T(z)| &= \left| \int_T^\infty f(t) e^{-zt}\, dt \right| \\
&\leq B \int_T^\infty e^{-\operatorname{Re}(z)t}\, dt = \frac{Be^{-\operatorname{Re}(z)T}}{\operatorname{Re} z} \\
\left|1 + \frac{z^2}{R^2}\right| &= \left|1 + \frac{z^2}{z\bar{z}}\right| = \left|\frac{z + \bar{z}}{\bar{z}}\right| = \frac{2 \operatorname{Re} z}{R}.
\end{aligned}$$

Also $\left|e^{zT}\right| = e^{\operatorname{Re}(z)T}$ and $\operatorname{length}(C_+) = \pi R$. Combining these inequalities,

$$|I_1| \leq \frac{B}{R}. \tag{90}$$

To estimate I_2, note that the integrand is an entire function of z, so the contour C_- can be replaced by the semicircle $C'_- = \{z \mid |z| = R \text{ and } \operatorname{Re} z < 0\}$. Again suppose that $|f(t)| \leq B$ for all t and observe that on C'_-, $|z| = R$ and

$$\begin{aligned}
|g_T(z)| &= \left| \int_0^T f(t) e^{-zt}\, dt \right| \\
&\leq B \int_{-\infty}^T e^{-\operatorname{Re}(z)t}\, dt = -\frac{Be^{-\operatorname{Re}(z)T}}{\operatorname{Re} z} \\
\left|1 + \frac{z^2}{R^2}\right| &= \left|1 + \frac{z^2}{z\bar{z}}\right| = \left|\frac{z + \bar{z}}{\bar{z}}\right| = -\frac{2 \operatorname{Re} z}{R}
\end{aligned}$$

Also $\left|e^{zT}\right| = e^{\mathrm{Re}\,(z)T}$ and length$(C'_-) = \pi R$. Combining these inequalities as with I_1 get

$$|I_2| \le \frac{B}{R}. \tag{91}$$

Finally to estimate I_3, note the the function $g(z)\left(1 + \frac{z^2}{R^2}\right)\frac{dz}{z}$ is continuous on C_- and therefore bounded there. Suppose it's absolute value less than or equal to M. Since $|e^{zT}| = e^{\mathrm{Re}\,(z)T} \le 1$ on C_- the integrand is also bounded in absolute value by M. Also $e^{zT} \to 0$ as $T \to \infty$ on C_- so the integrand converges to 0 as $T \to \infty$. By the bounded convergence theorem

$$\lim_{T\to\infty} I_3 = 0. \tag{92}$$

Combining (89), (90). (91), (92), get

$$\limsup_{T\to\infty} |g(0) - g_T(0)| \le \frac{2B}{R}$$

Since R is arbitrary, (87) follows. □

8.4. **Proof of PNT.** Now it is easy to complete the proof of PNT. The function $f(x) = \frac{\vartheta(x)}{x} - 1$ satisfies the hypothesis of Newman's Analytic Theorem II, since

$$\int_1^\infty \left(\frac{\vartheta(x)}{x} - 1\right) x^{-s}\, dx = -\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1} + h(s)$$

where $h(s)$ is holomorphic in $\mathrm{Re}\, s > 1/2$. Therefore

$$\int_1^\infty \left(\frac{\vartheta(x)}{x} - 1\right)\frac{dx}{x}$$

converges, and therefore as shown in (58), this proves PNT.

8.5. **PNT for arithmetic progressions.** The Prime Number Theorem states that

$$\pi(x) \sim \frac{x}{\log x} \qquad x \to \infty, \tag{93}$$

where $\pi(x) = \sum_{p\le x} 1$ is the prime counting function. Dirichlet's Theorem states that for an q and ℓ relatively prime to q the set

$$\{p\,|\,p \text{ prime and } p \equiv \ell \bmod q\} \tag{94}$$

is infinite. The Prime Number Theorem for arithmetic progressions says that the primes are evenly distributed between the $\varphi(q)$ sets of the form (94), that is, if the prime counting function of primes eqaul to ℓ mod q is defined by

$$\pi_\ell(x) = \sum_{\substack{p\equiv\ell\\ p\le x}} 1$$

then

$$\pi_\ell(x) \sim \frac{1}{\varphi(q)}\frac{x}{\log x} \qquad x \to \infty, \tag{95}$$

for every ℓ relatively prime to q.

The proof of (95) follows from Newman's Analytic Theorem in much the same way as PNT. Define the following functions

$$\begin{aligned} \vartheta_\ell(x) &= \sum_{\substack{p\equiv\ell\\ p\le x}} \log p \\ \Phi_\ell(s) &= \sum_{p\equiv\ell} \frac{\log p}{p^s} \\ &= s\int_1^\infty \vartheta_\ell(x)x^{-s-1}\,dx \end{aligned} \tag{96}$$

where the last integral follows for $\operatorname{Re} x > 1$ using summation by parts.

Using the same methods as in the proof of (48) it can be shown that as $x \to \infty$,

$$\pi_\ell(x) \sim \frac{1}{\varphi(q)}\frac{x}{\log x} \iff \vartheta_\ell(x) \sim \frac{1}{\varphi(q)}x. \tag{97}$$

Likewise arguing as in the proof of (54), for any character χ

$$\sum_p \chi(p)\frac{\log p}{p^s} + \frac{L'(s,\chi)}{L(s,\chi)} = h(s) \tag{98}$$

where $h(s)$ is a function holomorphic in $\operatorname{Re} s > 1/2$. By (98) and the orthogonality of characters

$$\begin{aligned} \varphi(q)\Phi_\ell(s) &= \varphi(q)\sum_{p\equiv\ell} p^{-s}\log p \\ &= \sum_\chi \overline{\chi(\ell)}\sum_p \chi(p)p^{-s}\log p \\ &= -\frac{\zeta'(s)}{\zeta(s)} - \sum_{\chi\ne\chi_0}\frac{L'(s,\chi)}{L(s,\chi)} + h_1(s) \end{aligned} \tag{99}$$

where $h_1(s)$ is a function holomorphic in $\operatorname{Re} s > 1/2$. Since the L functions for non-trivial characters are not zero at $s = 1$, (99) shows that $\varphi(q)\Phi_\ell(s)$ has a simple pole at $s = 1$ with residue 1.

The Prime Number Theorem for arithmetic progressions follows from the fact that for any character χ there are no zeros of $L(s,\chi)$ on $\operatorname{Re} s = 1$. We already showed this for the trivial character since this is equivalent to the fact that $\zeta(s)$ has no zeros on $\operatorname{Re} s = 1$. Now we show this for all L functions at once.

Theorem: For any character χ there are no zeros of $L(s,\chi)$ on $\operatorname{Re} s = 1$, and for any ℓ relatively prime to q no poles of $\Phi_\ell(s)$ except at $s = 1$.

Proof. This fact follows using the same method as in the proof that there are no zeros of $\zeta(s)$ on $\operatorname{Re} s = 1$. As in the proof of (85) we see that

$$\begin{aligned} 0 \le {}& \Phi_\ell(\sigma+2i\alpha) + 4\Phi_\ell(\sigma+i\alpha) + 6\Phi_\ell(\sigma) \\ &+ 4\Phi_\ell(\sigma-i\alpha) + \Phi_\ell(\sigma-2i\alpha). \end{aligned} \tag{100}$$

Suppose now that for $\alpha \ne 0$, the total number of zeros of $L(s,\chi)$ for all characters at $s = 1+i\alpha$ is μ and at $s = 1+i2\alpha$ is ν. Then by (99) $\varphi(q)\Phi_\ell$ has a simple pole with residue μ at $1+i\alpha$ and residue ν at $1+i2\alpha$. Since the coefficients of $\varphi(q)\Phi_\ell(s)$ are real, it also has a simple pole with residue μ at $s = 1-i\alpha$ and residue ν at $s = 1-i2\alpha$. Since for non trivial characters $L(1,\chi) \ne 0$, $\varphi(q)\Phi_\ell(s)$ has a simple pole with residue 1 at $s = 0$. Multiplying equation (100) by $\sigma - 1$ and letting $\sigma \to 1$ get $0 \le -2\nu - 8\mu + 6$. Since μ is a positive integer, conclude that $\mu = 0$. Thus there is no zero of any L function at $1 + i\alpha$. Since $\alpha \neq 0$ is arbitrary, there are no zeros of any L function on $\operatorname{Re} s = 1$. □

Now we have

The Prime Number Theorem for arithmetic progressions: For any ℓ relatively prime to q,

$$\vartheta_\ell(x) = \sum_{\substack{p \equiv \ell \\ p \le x}} \log p \sim \frac{x}{\varphi(q)}$$

Proof.

$$\int_1^\infty \left(\frac{\vartheta_\ell(x)}{x} - \frac{1}{\varphi(q)} \right) x^{-s}\, dx = \frac{\Phi_\ell(s)}{s} - \frac{1}{\varphi(q)} \frac{1}{s-1}.$$

for $\operatorname{Re} s > 1$ and by the previous theorem, this has analytic continuation to $\operatorname{Re} s \ge 1$. By the Chebychev estimate, $\vartheta_\ell(x) = \mathrm{O}(x)$ and the integrand is bounded, so Newman's Analytic Theorem can be applied and the integral

$$\int_1^\infty \left(\frac{\vartheta_\ell(x)}{x} - \frac{1}{\varphi(q)} \right) x^{-1}\, dx$$

converges. As with (58), since $\vartheta_\ell(x)$ is increasing, the convergence of the integral shows that $\vartheta_\ell(x)/x \sim 1/\varphi(q)$. □

8.6. **A counter example.** Here is an example to show that in the hypothesis of Newman's Theorem, it is necessary to have $g(z)$ be holomorphic on the entire line $\operatorname{Re} z = 0$. Let $f(t) = e^{i\alpha t}$ for $\alpha \neq 0$. Then for $\operatorname{Re} z > 0$,

$$g(z) = \int_0^\infty e^{i\alpha t} e^{-zt}\, dt = \frac{1}{z - i\alpha}$$

and g is holomorphic at 0 with $g(0) = -1/i\alpha$. However, for $z = 0$ the integral $\int_0^\infty e^{i\alpha t}\, dt$ does not converge. Here f is a complex valued function. For a real valued function take $f(t) = 2\cos \alpha t$.

In the second version of Newman's Theorem, take $f(x) = x^{i\alpha}$ for an example. This is the same counterexample as the one above, with $x = e^t$.

9. The Riemann Hypothesis

The zeta function is defined in $\operatorname{Re} s > 1$ by the Dirichlet series $\sum_{n=1}^\infty n^{-s}$ and the Euler product $\prod_p (1 - p^{-s})^{-1}$. To extend it to $\operatorname{Re} s > 0$ we use summation by parts as in (31) to write

$$\zeta(s) = \frac{1}{s-1} + 1 + s \int_1^\infty x^{-s-1}([x] - x)\, dx. \tag{101}$$

This show also that the logarithmic derivative ζ'/ζ extends to a meromorphic function in $\operatorname{Re} s > 0$ but we do not have a nice expression for it in $0 < \operatorname{Re} s \le 1$. We do know that the difference $\displaystyle\sum_p \frac{\log p}{p^s} - \frac{\zeta'(s)}{\zeta(s)}$ extends to a function holomorphic in $\operatorname{Re} s > 1/2$.

The Prime Number Theorem is equivalent to the fact that $\zeta(s)$ has no zeros in $\operatorname{Re} s \ge 1$. The *Riemann Hypothesis* (RH) states the stronger conclusion that

$\zeta(s)$ has no zeros in $\operatorname{Re} s > 1/2$. This conjecture has not yet been proved. We will explain how RH relates to the distribution of prime numbers. First we discuss alternate formulations of PNT in terms of sums or averages of coefficients of various Dirichlet series.

9.1. **PNT and $M(x)$.** Recall that Mertens function $M(x) = \sum_{n \le x} \mu(n)$ is given by the partial sums of the Möbius function defined by (61). The Möbius function is defined by the Dirichlet series

$$\frac{1}{\zeta(s)} = \sum_n \mu(n) n^{-s} \tag{102}$$

obtained from the Euler factorization of ζ. Since $|\mu(n)| = 1$ for all n, $|M(x)| \le x$, and partial summation gives

$$\frac{1}{\zeta(s)} = s \int_1^\infty M(x) x^{-s-1}\, dx \tag{103}$$

for $\operatorname{Re} s > 1$.

The Prime Number Theorem is equivalent to the fact that the coefficients of $1/\zeta$ average to zero. This can be interpreted as saying that the sequence $\mu(n)$ behaves like a random sequence of integers -1, 0, 1.

$$\text{PNT} \iff \frac{M(x)}{x} \to 0 \text{ as } x \to \infty. \tag{104}$$

Proof. First show PNT $\implies M(x)/x \to 0$.

Recall that PNT is equivalent to the fact that $\zeta(s)$ has no zeros on $\operatorname{Re} s = 1$, The function $M(x)/x$ is clearly bounded. Using Newman's analytic theorem and (103) we find that if $\zeta(s)$ has no zeros on $\operatorname{Re} s = 1$, then

$$\int_1^\infty \frac{M(x)}{x} \frac{dx}{x} = 0 \tag{105}$$

since by Newman's theorem and (103) the integral converges to $1/\zeta(s)|_{s=1} = 0$. The convergence of the integral implies that $M(x)/x \to 0$ as $x \to \infty$. To see this note that for $t \ge x$, $M(x) - M(t) = \displaystyle\sum_{x<n\le t} \mu(n)$. Since $-1 \le \mu(n) \le 1$,

$$x - t \le M(x) - M(t) \le t - x \tag{106}$$

Now given $\epsilon > 0$ suppose that

$$M(x) \ge \epsilon x. \tag{107}$$

Then for $x \le t \le \left(1 + \frac{\epsilon}{2}\right) x$ we have by (106)

$$M(t) \ge M(x) + x - t \ge \frac{\epsilon}{2} x.$$

So

$$\int_x^{(1+\frac{\epsilon}{2})x} \frac{M(t)}{t} \frac{dt}{t} \ge \frac{\epsilon}{2} \int_x^{(1+\frac{\epsilon}{2})x} \frac{x}{t^2}\, dt = \frac{\epsilon}{2} \int_1^{(1+\frac{\epsilon}{2})} \frac{du}{u^2} > 0.$$

But the right hand side is independent of x so, if (107) is true for a sequence x going to infinity, then the integral does not converge.

A similar argument holds for

$$M(x) \le -\epsilon x. \tag{108}$$

So $M(x)/x \to 0$.

Next show $M(x)/x \to 0 \implies$ PNT. By (103) for $\operatorname{Re} s > 1$,

$$\frac{1}{s\zeta(s)} = \int_1^\infty \frac{M(x)}{x} x^{-s}\, dx. \tag{109}$$

Given $\epsilon > 0$ suppose there is an $N > 0$ such that $|M(x)/x| \le \epsilon$ for $x \ge N$. Then similar to (81) we have for $s = \sigma + i\tau$, $\sigma > 1$,

$$\left| \frac{1}{s\zeta(s)} \right| \le \int_1^N x^{-\sigma}\, dx + \epsilon \frac{N^{-\sigma+1}}{\sigma - 1},$$

so

$$\left| \lim_{\sigma \to 1^+} (\sigma - 1) \frac{1}{s\zeta(s)} \right| \le \epsilon.$$

Since this is true for all ϵ,

$$\lim_{\sigma \to 1^+} (\sigma - 1) \frac{1}{s\zeta(s)} = 0$$

and so there are no singularities of $1/\zeta$ on the line $\operatorname{Re} s = 1$.

□

9.2. **PNT and $1/\zeta$.** For $\operatorname{Re} s > 1$, the Euler factorization gives $1/\zeta(s) = \sum_n \mu(n) n^{-s}$. For the analytic continuation of ζ we easily see that since $\zeta(s)$ has a pole at $s = 1$ it follows that $\frac{1}{\zeta}(1) = 0$. It might be conjectured that the series converges at $s = 1$ to 0. In fact, this is equivalent to PNT,

$$\text{PNT} \iff \sum_{n=1}^\infty \frac{\mu(n)}{n} = 0. \tag{110}$$

A direct proof can be found in [1] p. 97. Here is another proof.

Proof. First show PNT $\implies$ (110). Let $A(x) = \sum_{n \le x} \frac{\mu(n)}{n}$. Note that

$$|A(x)| \le \sum_{n \le x} \frac{1}{n} = \mathrm{O}(\log x) \tag{111}$$

Using summation by parts. for $\operatorname{Re} s > 1$

$$\begin{aligned} A(x) &= \int_1^x \frac{dM(t)}{t} \\ &= \frac{M(x)}{x} + \int_1^x \frac{M(t)}{t^2}\, dt. \end{aligned}$$

Now by (111) the first term converges to 0 as $x \to \infty$, and by (105) the second term converges to 0 as $x \to \infty$.

Next show (110) $\implies$ PNT. By (102) we get using summation by parts

$$\begin{aligned} \frac{1}{\zeta(s)} &= \sum_n \frac{\mu(n)}{n} n^{-s+1} = \int_1^\infty x^{-s+1} dA(x) \\ &= (s-1) \int_1^\infty A(x) x^{-s}\, dx \end{aligned} \tag{112}$$

So

$$\frac{1}{\zeta(s)(s-1)} = \int_1^\infty A(x) x^{-s}\, dx. \tag{113}$$

Let $\epsilon > 0$ be given and suppose $|A(x)| < \epsilon$ for $x > N$. Then if $\sigma = \operatorname{Re} s$,

$$\begin{aligned} \left| \frac{1}{\zeta(s)(s-1)} \right| &\leq \epsilon \int_M^\infty x^{-\sigma+1}\, dx + g(\sigma) \\ &= \epsilon \frac{M^{-\sigma+1}}{\sigma - 1} + g(\sigma) \end{aligned} \tag{114}$$

where g is a continuous function. This shows

$$\left| \lim_{\sigma \to 1^+} \frac{1}{\zeta(s)(s-1)} \right| < \epsilon$$

for every $\epsilon > 0$ and hence the limit is 0 and there are no zeros of $\zeta(s)$ on the line $\operatorname{Re} s = 1$.

□

This is a Tauberian type result and one of the results that Chebychev needed to complete his attempted proof of PNT in section 7 since his argument can be summarized as

$$\begin{aligned} \psi(x) &= \sum_n \mu(n) T(x/n) \\ &= \sum_n \mu(n) \left(\frac{x}{n} \log\left(\frac{x}{n}\right) - \frac{x}{n} \right) \\ &= (x \log x - x) \left(\sum_n \frac{\mu(n)}{n} \right) - x \left(\sum_n \mu(n) \frac{\log n}{n} \right). \end{aligned} \tag{115}$$

This reduces to x if

$$\sum_n \frac{\mu(n)}{n} = 0 \text{ and } \sum_n \mu(n) \frac{\log n}{n} = -1.$$

The first equation we showed is equivalent to PNT. We now discuss the second.

9.3. **PNT and $1/\zeta'$.** Similarly, for $\operatorname{Re} s > 1$, $(1/\zeta(s))' = -\sum_n \mu(n) n^{-s} \log n$. For the analytic continuation of ζ since $\zeta(s)$ has a pole at $s = 1$ with residue 1 we see that $(\frac{1}{\zeta})'(1) = 1$. It might be conjectured that the series converges at $s = 1$. In fact

$$\text{PNT} \iff \sum_{n=1}^\infty \mu(n) \frac{\log n}{n} = -1. \tag{116}$$

(I think this is true and I think I saw it in a book, but I can't prove it. It might require a stronger form of Newman's Theorem.) This is a Tauberian type result and is the other fact that Chebychev needed to complete his attempted proof of PNT in section 7.

9.4. **PNT and M_1.** The Möbius function $\mu(n)$ takes values -1, 0, 1. A similar function which takes values ± 1 is the *Liouville function* $\lambda(n)$. This function is defined as the coefficients of the Dirichlet series for $\zeta(2s)/\zeta(s)$. From the Euler product for ζ,

$$\begin{aligned}\frac{\zeta(2s)}{\zeta(s)} &= \prod_p \left(\frac{1-p^{-2s}}{1-p^{-s}}\right)^{-1} \\ &= \prod_p (1+p^{-s})^{-1} \\ &= \prod_p \sum_k (-1)^k p^{-ks} \\ &= \sum_n \lambda(n) n^{-s}.\end{aligned} \tag{117}$$

From this we see that

$$\lambda(n) = (-1)^{a_1+\cdots+a_k} \text{ if } n \text{ factors into primes as } n = p_1^{a_1}\cdots p_k^{a_k}. \tag{118}$$

This is a variation of the Möbius function. Define the partial sum by

$$M_1(x) = \sum_{n\leq x} \lambda(n). \tag{119}$$

Clearly $M_1(x) \leq x$ so by partial summation,

$$\frac{\zeta(2s)}{\zeta(s)} = s\int_1^\infty M_1(x)x^{-s-1}\,dx. \tag{120}$$

Now the same argument as in section 9.1 shows that

$$\text{PNT} \iff \frac{M_1(x)}{x} \to 0 \text{ as } x \to \infty. \tag{121}$$

Again this says that the sequence $\lambda(n)$ behaves like a random sequence of integers ± 1.

9.5. **RH and μ and λ.** It is easily seen that

$$\left\{ \text{For all } \epsilon > 0,\ \lim_{x\to\infty} \frac{|M(x)|}{x^{\frac{1}{2}+\epsilon}} = 0 \right\} \implies \text{RH} \tag{122}$$

since by (103) the integral converges absolutely and uniformly for $\operatorname{Re} s > \frac{1}{2}+\epsilon$ and so $1/\zeta(s)$ is analytic there and $\zeta(s)$ has no zeros there. Since this is true for every ϵ, $\zeta(s)$ has no zeros in $\operatorname{Re} s > 1/2$. Actually, the statement on the left in (122) is equivalent to RH, but the implication in the other direction is harder to prove.

Similarly by (120),

$$\left\{ \text{For all } \epsilon > 0,\ \lim_{x\to\infty} \frac{|M_1(x)|}{x^{\frac{1}{2}+\epsilon}} = 0 \right\} \implies \text{RH} \tag{123}$$

Again, the two sides are equivalent, but the implication in the opposite direction is harder to prove. The statement on the left is what we would expect of a random sequence of ± 1, as we show in the next section.

9.6. **RH is almost certainly true.** Although there is some reasoning behind this statement, it really makes no sense. But reasons for stating that RH is almost certainly true give some insight into the meaning of RH.

We think of the sequence $\lambda(n)$ as a random sequence of numbers ± 1 and ask what is the probability that (123) holds. We show that the probability is 1. To see this, first consider the question when we have a finite sequence of length N.

Consider the situation where $\lambda(n)$ is a random choice of a sequence of N values ± 1. There are 2^N such sequences. If the sequence consists of k values 1 and $N-k$ values -1 then the sum S is $2k-N$. There are $\binom{N}{k}$ such sequences. Thinking of S as a random variable, the probability that $S = 2k-N$ is $2^{-N}\binom{N}{k}$, and the distribution of S is the binomial distribution

$$f_N(x) = \frac{1}{2^N}\sum_{k=0}^{N}\binom{N}{k}\delta_{2k-N}(x), \tag{124}$$

where δ is the Dirac delta function. The mean of f_N is zero and the variance is

$$\int_{-\infty}^{\infty} f_N(x)x^2\,dx = \frac{1}{2^N}\sum_{k=0}^{N}\binom{N}{k}(2k-N)^2 = N.$$

Comparing to (123), look at the probability that $|S| \le N^{\frac{1}{2}+\epsilon}$,

$$P(|S| \le N^{\frac{1}{2}+\epsilon}) = \frac{1}{2^N}\sum_{|2k-N|\le N^{\frac{1}{2}+\epsilon}}\binom{N}{k}. \tag{125}$$

We show that

$$P(|S| \le N^{\frac{1}{2}+\epsilon}) \to 1, \text{ as } N \to \infty. \tag{126}$$

This means that for almost all choices of random sequences λ with values ± 1, (123) holds. The idea of the proof of (126) is that the binomial distribution converges to the Gaussian distribution. More precisely, if the binomial distribution (124) is scaled by $\sqrt{N}$ to $\sqrt{N}f_N(\sqrt{N}x)$, which is a probability distribution with mean 0 and variance 1, then it converges as $N \to \infty$ to the Gaussian distribution $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ with the same mean and variance. So now the proof can be summarized as

$$\begin{aligned} P(|S| \le N^{\frac{1}{2}+\epsilon}) &= \int_{-N^{\frac{1}{2}+\epsilon}}^{N^{\frac{1}{2}+\epsilon}} f_N(x)\,dx \\ &= \int_{-N^{\epsilon}}^{N^{\epsilon}} \sqrt{N}f_N(\sqrt{N}x)\,dx \\ &\sim \frac{1}{\sqrt{2\pi}}\int_{-N^{\epsilon}}^{N^{\epsilon}} e^{-\frac{x^2}{2}}\,dx \\ &\to 1 \text{ as } N \to \infty. \end{aligned} \tag{127}$$

10. Faulhaber and Bernoulli polynomials

10.1. **Faulhaber polynomials.** Johann Faulhaber (1580 – 1635) is a German mathematician famous for his discovery of polynomials giving sums of integer powers of integers. Examples for $\mu = 0, \ldots, 4$ are:

$$\sum_{n=1}^{N} 1 = N$$

$$\sum_{n=1}^{N} n = \frac{1}{2}N^2 + \frac{1}{2}N$$

$$\sum_{n=1}^{N} n^2 = \frac{1}{3}N^3 + \frac{1}{2}N^2 + \frac{1}{6}N$$

$$\sum_{n=1}^{N} n^3 = \frac{1}{4}N^4 + \frac{1}{2}N^3 + \frac{1}{4}N^2$$

$$\sum_{n=1}^{N} n^4 = \frac{1}{5}N^5 + \frac{1}{2}N^4 + \frac{1}{3}N^3 - \frac{1}{30}N$$

As a function of $s = -\mu$ we can think of these as similar to the ζ function, but finite sums.

These polynomials are easy to find by induction. Let

$$S(\mu, N) = \sum_{n=1}^{N} n^{\mu} \tag{128}$$

denote the sum of μth powers, where μ is a non-negative integer. Then

$$\begin{aligned}(N+1)^{\mu+1} - 1 &= \sum_{n=1}^{N} \left((n+1)^{\mu+1} - n^{\mu+1}\right) \\ &= \sum_{n=1}^{N}\sum_{k=0}^{\mu} \binom{\mu+1}{k} n^k \\ &= \sum_{k=0}^{\mu} \binom{\mu+1}{k} \sum_{n=1}^{N} n^k \\ &= \sum_{k=0}^{\mu} \binom{\mu+1}{k} S(k, N)\end{aligned} \tag{129}$$

so from $S(k, N)$, $k = 1, \ldots, \mu - 1$ equation (129) gives $S(\mu, N)$. For example for $\mu = 2$ and $S(0, N) = N$ and $S(1, N) = N(N + 1)/2$, (129) gives

$$(N+1)^3 - 1 = N + 3N(N+1)/2 + S(2, N)$$

which can be solved to give

$$S(2, N) = \frac{1}{3}N^3 + \frac{1}{2}N^2 + \frac{1}{6}N.$$

This method to generate the Faulhaber polynomials by induction is due to Pascal (1864).

10.2. **Bernoulli polynomials.** The Bernoulli polynomials are almost the same as the Faulhaber polynomials but Bernoulli polynomials are generated as coefficients of a power series. The Bernoulli numbers are defined by the generating function

$$\frac{1}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} t^{k-1}. \tag{130}$$

The coefficients B_k can be computed from the geometric series

$$\begin{aligned} \frac{t}{e^t - 1} &= \frac{1}{1 + \frac{1}{2}t + \frac{1}{6}t^2 + \cdots} \\ &= 1 - \left(\frac{1}{2}t + \frac{1}{6}t^2 + \cdots\right) + \left(\frac{1}{2}t + \frac{1}{6}t^2 + \cdots\right)^2 + \cdots \end{aligned}$$

The first few Bernoulli numbers are

$$B_0 = 1 \quad B_1 = -\frac{1}{2} \quad B_2 = \frac{1}{6} \quad B_4 = -\frac{1}{30} \quad B_6 = \frac{1}{42}.$$

Since

$$\frac{1}{e^t - 1} - \frac{1}{t} + \frac{1}{2}$$

is an odd function, the Bernoulli numbers B_k for k odd, $k \geq 3$ are zero.

The Bernoulli polynomials are defined by the generating function

$$\frac{e^{xt}}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k(x)}{k!} t^{k-1}. \tag{131}$$

Using

$$e^{xt} = \sum_{k=0}^{\infty} \frac{x^k}{k!} t^k$$

and multiplying by the series (130) it follows from (131) that the Bernoulli polynomials are given from the Bernoulli numbers by

$$B_\mu(x) = \sum_{k=0}^{\mu} \binom{\mu}{k} B_k x^{\mu-k}. \tag{132}$$

Using the *umbral calculus* where B_k is replaced by B^k, (132) can be written as

$$B_\mu(x) = (B + x)^\mu. \tag{133}$$

This is a good mnemonic device for remembering the formula, but the umbral calculus can be given a rigorous foundation.

The first few Bernoulli polynomials are

$$B_0(x) = 1 \qquad B_1(x) = x - \frac{1}{2} \qquad B_2(x) = x^2 - x + \frac{1}{6}$$

$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x \qquad B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30}$$

Note that

$$B_k = B_k(0). \tag{134}$$

so the Bernoulli numbers are the constant terms in the Bernoulli polynomials.

10.3. **The relationship between Bernoulli polynomials and Faulhaber polynomials.** The Bernoulli polynomials are almost the same as the Faulhaber polynomials ([3]). Write

$$\begin{aligned}\frac{e^{(N+1)t}-1}{e^t-1} &= \sum_{n=0}^{N} e^{nt} \\ &= \sum_{n=0}^{N}\sum_{\mu=0}^{\infty}\frac{1}{\mu!}n^\mu t^\mu \\ &= \sum_{\mu=0}^{\infty}\frac{1}{\mu!}S(\mu,N)t^\mu\end{aligned} \tag{135}$$

where $S(\mu, N)$ denotes the Faulhaber sum $S(\mu, N) = \sum_{n=0}^{N} n^\mu$. (Note, for $\mu = 0$ define $0^0 = 1$.) Now expand the left hand side of (135) in powers of t using (131):

$$\frac{e^{(N+1)t}}{e^t-1} = -\frac{e^{(-N)(-t)}}{e^{-t}-1} = \sum_{\mu=0}^{\infty}(-1)^{\mu+1}\frac{B_\mu(-N)}{\mu!}t^{\mu-1}, \tag{136}$$

so using (130)

$$\frac{e^{(N+1)t}-1}{e^t-1} = \sum_{\mu=0}^{\infty}\frac{(-1)^{\mu+1}B_\mu(-N)-B_\mu}{\mu!}t^{\mu+1}. \tag{137}$$

Now equating coefficients of t^μ in (137) and (135), get

$$\begin{aligned}\sum_{n=0}^{N} n^\mu &= \frac{1}{\mu+1}\left[(-1)^{\mu+1}B_{\mu+1}(-N) - B_{\mu+1}\right] \\ &= \frac{1}{\mu+1}\sum_{k=0}^{\mu}(-1)^k\binom{\mu+1}{k}B_k N^{\mu+1-k}.\end{aligned} \tag{138}$$

and this shows that Faulhaber polynomials are easily written in terms of Bernoulli polynomials.

Using the umbral calculus,

$$\sum_{n=0}^{N} n^\mu = \frac{1}{\mu+1}\left((N-B)^{\mu+1} - B^{\mu+1}\right) \tag{139}$$

(Note: Since all the odd Bernoulli numbers except B_1 are zero, if B_1 is defined to be $1/2$ instead of $-1/2$ then the above could be written $\frac{1}{\mu+1}\left((B+N)^{\mu+1} - B^{\mu+1}\right)$).

Similar techniques using (135) give

$$\sum_{n=0}^{N} n^\mu = \frac{1}{\mu+1}\left((B+N+1)^{\mu+1} - B^{\mu+1}\right) \tag{140}$$

or using more umbral calculus,

$$\sum_{n=0}^{N} n^\mu = \int_0^{N+1}(B+x)^\mu\,dx. \tag{141}$$

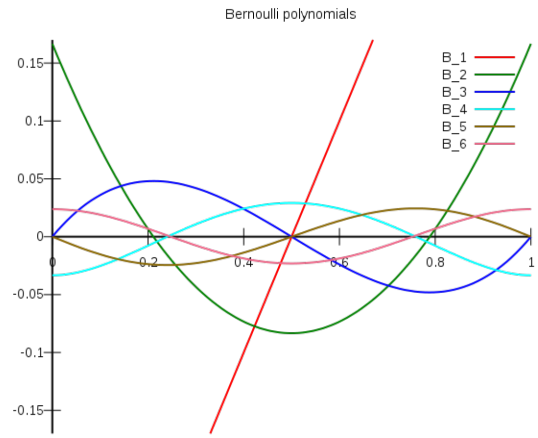

FIGURE 1. Bernoulli Polynomials

10.4. **Properties of Bernoulli polynomials.** From (131) we see that for $k \geq 1$,

$$\int_0^1 B_k(x)\,dx = 0 \quad k \geq 1 \tag{142}$$

$$B_k(x+1) = (-1)^k B_k(-x) \tag{143}$$

$$B_k(x+1) - B_k(x) = kx^{k-1} \tag{144}$$

$$B_k'(x) = kB_{k-1}(x). \tag{145}$$

To prove (142), for example, integrate both sides of (131) from 0 to 1 and use the fact that

$$\int_0^1 \frac{e^{xt}}{e^t - 1}\,dx = \frac{1}{t}.$$

10.5. **Periodic Bernoulli functions.** From (144) it follows that

$$B_k(0) = B_k(1) \qquad k \geq 2 \tag{146}$$

The Bernoulli polynomials restricted to the interval $[0,1]$ can be extended to a function

$$P_k(x) = B_k(x - [x]) \tag{147}$$

on $\mathbb{R}$ of period 1. By (146), the functions P_k are continuous for $k \geq 2$. Also

$$P_k(n) = B_k \qquad k \geq 2 \text{ and } n \text{ an integer.} \tag{148}$$

It follows from (145) that

$$P_k'(x) = kP_{k-1}(x) \text{ except when } x \text{ is an integer.} \tag{149}$$

Also from (142),

$$\int_M^N P_k(x)\,dx = 0 \qquad \text{for } M \text{ and } N \text{ integers and } k \geq 1. \tag{150}$$

11. Euler-MacLaurin Summation

The Euler-Maclaurin summation formula is a method to generalize the method of summation by parts for $\sum_{n=M}^{N} f(n)$ where f is a differentiable function (23). The sum is converted into an integral to make the process look like integration by parts. If f has more derivatives, the integration by parts can be continued giving a more accurate approximation for the sum. The Bernoulli polynomials play an important role.

We might ask what would have happened if Faulhaber investigated the sum of fractional powers of integers. It will follow from Euler-Mclaurin summation that he would have a similar formula but with a remainder. That remainder is related to the the zeta function.

11.1. **The Euler-Mclaurin summation formula.**

Summation by parts gives a way to express the difference between the sum and the corresponding integral using the notation of integration by parts:

$$\begin{aligned}\sum_{n=M}^{N} f(n) - \int_{M}^{N} f(x)\,dx &= \int_{M^-}^{N} f(x)\,d\left([x]-x\right) \\ &= f(x)([x]-x)\Big|_{M^-}^{N} + \int_{M}^{N} f'(x)(x-[x])\,dx \\ &= f(M) + \int_{M}^{N} f'(x)(x-[x])\,dx.\end{aligned} \tag{151}$$

This technique is slightly modified by replacing $x-[x]$ by $x-[x]+1/2$:

$$\begin{aligned}\sum_{n=M}^{N} f(n) - \int_{M}^{N} f(x)\,dx &= \int_{M^-}^{N} f(x)\,d\left([x]-x+1/2\right) \\ &= f(x)([x]-x+1/2)\Big|_{M^-}^{N} + \int_{M}^{N} f'(x)(x-[x]+1/2)\,dx \\ &= \frac{1}{2}(f(M)+f(N)) + \int_{M}^{N} f'(x)(x-[x]+1/2)\,dx.\end{aligned} \tag{152}$$

Since the first Bernoulli polynomial is $B_1(x) = x - 1/2$, equation (152) can be written as

$$\begin{aligned}&\sum_{n=M}^{N} f(n) - \int_{M}^{N} f(x)\,dx \\ &\qquad = \frac{1}{2}(f(M)+f(N)) + \int_{M}^{N} f'(x)B_1(x-[x])\,dx.\end{aligned} \tag{153}$$

This integration by parts can be done again. Using (145) write $B_1(x-[x])\,dx = dB_2(x-[x])/2$, or equivalently $P_1(x)\,dx = P_2(x)/2$ where P_k are the periodic Bernoulli functions described in the previous section. Now integrate (153) again by parts giving

$$\begin{aligned}\sum_{n=M}^{N} f(n) - \int_M^N f(x)\,dx \\ &= \frac{1}{2}(f(M)+f(N)) + \frac{B_2}{2} f'(x)\big|_M^N \\ &- \frac{1}{2}\int_M^N f''(x)B_2(x-[x])\,dx.\end{aligned} \tag{154}$$

Continuing to integrate by parts, obtain the *Euler-Mclaurin summation formula*

$$\begin{aligned}\sum_{n=M}^{N} f(n) - \int_M^N f(x)\,dx \\ &= \frac{1}{2}(f(M)+f(N)) + \sum_{k=2}^{K} \frac{(-1)^k B_k}{k!} f^{(k-1)}(x)\bigg|_M^N \\ &+ R_K\end{aligned} \tag{155}$$

where the remainder R_K is given by

$$R_K = (-1)^{K+1}\frac{1}{K!}\int_M^N f^{(K)}(x)B_K(x-[x])\,dx. \tag{156}$$

Note that the $(-1)^k$ is unnecessary in the sum since the odd Bernoulli numbers are 0 for $k \geq 2$.

11.2. **Faulhaber sums from Euler-Mclaurin.** The Faulhaber/Bernoulli polynomials for sums of integer powers of integers can be obtained from the Euler-Mclaurin formula (155). For an integer $\mu \geq 1$ let $f(x) = x^\mu$ and let $M = 0$. Since the μ derivative of x^μ is $\mu!$, the μ remainder term is

$$R_\mu = \mu! \int_0^N B_\mu(x-[x])\,dx = 0 \tag{157}$$

by (142). So now (155) becomes

$$\begin{aligned}\sum_{n=0}^{N} n^\mu &= \frac{N^{\mu+1}}{\mu+1} + \frac{1}{2}N^\mu \\ &+ \sum_{k=2}^{\mu} \frac{(-1)^k B_k}{k!}\mu(\mu-1)\cdots(\mu-k+2)N^{\mu-k+1}\end{aligned} \tag{158}$$

which is the same as (138). Note: if $\mu = 1$ take just the first two terms on the right.

11.3. **Euler-Mclaurin and the zeta function.** The Euler-Mclaurin formula (155) can be applied to any complex valued function. Taking $f(x) = x^{-s}$ where s is any complex number, and taking $M = 1$, (155) gives for $s \neq 1$,

$$\begin{aligned}\sum_{n=1}^{N} n^{-s} &= \frac{N^{-s+1}}{-s+1} + \frac{1}{2}N^{-s} + c(s) \\ &+ \sum_{k=2}^{K} \frac{(-1)^k B_k}{k!}(-s)(-s-1)\cdots(-s-k+2)N^{-s+1-k} + R_K\end{aligned} \tag{159}$$

where

$$R_K = (-1)^{K+1}\frac{(-s)(-s-1)\cdots(-s-K+1)}{K!}\int_1^N x^{-s-K}B_K(x-[x])\,dx \tag{160}$$

and where

$$c(s) = \frac{1}{s-1} - \frac{1}{2} - \sum_{k=2}^{K}\frac{(-1)^k B_k}{k!}(-s)(-s-1)\cdots(-s-k+2). \tag{161}$$

Note that $c(s)$ is holomorphic except for a pole at $s = 1$.

Equation (159) can be written more succinctly as

$$\sum_{n=1}^{N} n^{-s} = \frac{1}{-s+1}\sum_{k=0}^{K}\binom{-s+1}{k}(-1)^k B_k N^{-s+1-k} + c(s) + R_K \tag{162}$$

with

$$R_K = (-1)^{K+1}\binom{-s}{K}\int_1^N x^{-s-K}B_K(x-[x])\,dx \tag{163}$$

where for z a complex number and k a non-negative integer, we define

$$\binom{z}{k} = \frac{z(z-1)\cdots(z-k+1)}{k!}.$$

If z is a positive integer and $k \le z$, this is the binomial coefficient. Equation (162) is similar to Faulhaber's formula (138). In (162) we could possibly take $K = 1$.

Now suppose that $\operatorname{Re} s > -K+1$ and look at the remainder (163). The function $B_K(x-[x])$ is bounded and so the absolute value of the integrand is $\le Cx^{-\operatorname{Re} s-K}$ where C is a positive constant. It follows that the integral for R_K converges as $N \to \infty$ to a holomorphic function $f(s)$ and that $R_K = f(s) + \mathrm{O}(N^{-\operatorname{Re} s-K+1})$. Now it follows that in $\operatorname{Re} s > -K+1$,

$$\begin{aligned}\sum_{n=1}^{N} n^{-s} = \frac{1}{-s+1}\sum_{k=0}^{K}\binom{-s+1}{k}(-1)^k B_k N^{-s+1-k}\\ + \zeta(s) + \mathrm{O}(N^{-\operatorname{Re} s-K+1})\end{aligned} \tag{164}$$

where $\zeta(s)$ is holomorphic except for a pole at $s = 1$. If $\operatorname{Re} s > 1$, all terms on the right hand side of (164) go to zero as $N \to \infty$. So for $\operatorname{Re} s > 1$, $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ and ζ is the analytic continuation of the Riemann zeta function.

References

[1] Tom M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. Springer, 1976.

[2] P.T. Bateman and H.G. Diamond. A hundred years of prime numbers. *The American mathematical monthly*, 103(9):729–741, 1996.

[3] John Conway and Richard Guy. *The Book of Numbers*. Springer, 1995.

[4] H. M. Edwards. *Riemann's Zeta Function*. Dover, 2001.

[5] Donald J. Newman. *Analytic Number Theory*, volume 177 of *Graduate Texts in Mathematics*. Springer, 1998.

[6] Elias M. Stein and Rami Shakarchi. *Complex Analysis*, volume 2 of *Princeton Lectures in Analysis*. Princeton University Press, 2003.

[7] Elias M. Stein and Rami Shakarchi. *Fourier Analysis: An Introduction*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, 2003.

[8] D Zagier. Newman's short proof of the prime number theorem. *American Mathematical Monthly*, Dec 1997.