THE ZYGMUND MORSE-SARD THEOREM

Alec Norton

Abstract. The classical Morse-Sard Theorem says that the set of critical values of $f : \mathbf{R}^{n+k} \to \mathbf{R}^n$ has Lebesgue measure zero if $f \in C^{k+1}$. We show the C^{k+1} smoothness requirement can be weakened to $C^{k+\text{Zygmund}}$. This is corollary to the following theorem: For integers $n > m > r > 0$, let $s = (n-r)/(m-r)$; if $f : \mathbf{R}^n \to \mathbf{R}^m$ belongs to the Lipschitz class Λ_s and E is a set of rank r for f, then $f(E)$ has measure zero.

0. Introduction

Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a differentiable function. Must the set of critical values of f have measure zero? The answer is "yes" provided that f is sufficiently smooth, and the classical theorem in this regard is the Morse-Sard theorem (often called simply "Sard's theorem" — A.P. Morse [6] proved the theorem in 1939 for the real-valued case; A. Sard [10] then extended that result to the vector-valued case.)

To state the theorem, we need some terminology. For f as above, a point $x \in \mathbf{R}^n$ is called a *critical point* if the linear mapping $Df(x)$ is not surjective; a *critical value* is the image under f of a critical point. The set of all critical values is a subset of the target space $\mathbf{R}^m$.

The Morse-Sard Theorem. *[6,10]. Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be of class C^k.*

If $k \geq \max\{n-m+1, 1\}$, then the set of critical values of f has Lebesgue measure zero.

We henceforth restrict our attention to the case $n > m$; smoothness is not an issue when $n \leq m$. (In fact even measurability is not required if $n \leq m$. See Varberg [12].)

Prior to Morse's work, H. Whitney had established in a famous paper [13] that *some* differentiability requirement is necessary by constructing an example of a C^1 function $f : \mathbf{R}^2 \to \mathbf{R}$ not constant on a connected set of critical points (and hence having a nontrivial interval of critical values). This example, and related ones in higher dimensions, shows that the smoothness hypothesis of the Morse-Sard theorem is sharp to the extent that one cannot replace the integer $n-m+1$ by any smaller integer. In fact, more is true: for every n, m with $n > m > 0$ there is a function $f \in \bigcap_{\alpha<1} \mathbf{C}^{n-m,\alpha}(\mathbf{R}^n, \mathbf{R}^m)$ whose critical value set contains an open set [2,8].

(Notation: for $\alpha \in [0, 1)$, k a nonnegative integer, we say $f \in C^{k,\alpha}(\mathbf{R}^n, \mathbf{R}^m)$ if $f : \mathbf{R}^n \to \mathbf{R}^m$ is of class C^k and the kth derivative $D^k f$ locally satisfies a Hölder condition with exponent α. We say $f \in C^{k,1}(\mathbf{R}^n, \mathbf{R}^m)$ if the kth derivative is locally Lipschitz.)

Typeset by $\mathcal{AMS}$-TEX

This paper addresses the question of whether the Morse-Sard theorem holds true for any natural smoothness class weaker than C^{n-m+1}. Because of the examples mentioned above, there is very little room for improvement. Nevertheless, S. Bates [1] has recently improved the statement to

The Lipschitz Morse-Sard Theorem (Bates). *Let n, m be positive integers with $n > m$. If $f \in C^{n-m,1}(\mathbf{R}^n, \mathbf{R}^m)$ then the set of critical values of f has Lebesgue measure zero.*

Even weaker than the Lipschitz class is the so-called *Zygmund class*, defined as follows. We say that $f : \mathbf{R}^n \to \mathbf{R}^m$ is *Zygmund* if it is continuous and for every compact set $K \subset \mathbf{R}^n$, there is a constant C such that if $x, x+h$, and $x-h$ belong to K, then

$$|f(x+h) + f(x-h) - 2f(x)| \leq C|h|.$$

We say that $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R}^m)$ if f is C^k and all of the kth partial derivatives of f are Zygmund.

For basic properties of the Zygmund class, see [5] or [11]. This class is very well known in harmonic analysis, and in fact there is a strong case (see [5]) that for many purposes in analysis the class $C^{k,Z}$ is more natural than the classical class C^{k+1}. This class also arises naturally in certain dynamical systems settings, e.g. [9].

It will be slightly more convenient for us to use the "Lipschitz spaces" Λ_s of harmonic analysis, defined as follows.

If $0 < s < 1$, then $f \in \Lambda_s(\mathbf{R}^n, \mathbf{R}^m)$ if $f : \mathbf{R}^n \to \mathbf{R}^m$ is continuous and

$$||f||_s \equiv \sup_{x \in \mathbf{R}^n} |f(x)| + \sup_{x,h \in \mathbf{R}^n} |f(x+h) - f(x)|/|h|^s < \infty.$$

We say $f \in \Lambda_1$ if f is continuous and

$$||f||_1 \equiv \sup_{x \in \mathbf{R}^n} |f(x)| + \sup_{x,h \in \mathbf{R}^n} |f(x+h) + f(x-h) - 2f(x)|/|h| < \infty.$$

Inductively, for $s > 1$, we say $f \in \Lambda_s(\mathbf{R}^n, \mathbf{R}^m)$ if $f \in C^k(\mathbf{R}^n, \mathbf{R}^m)$, where k is the largest integer less than s, and

$$||f||_s \equiv ||f||_{s-1} + \sum_{j=1}^{n} ||\frac{\partial f}{\partial x_j}||_{s-1} < \infty.$$

Note that the space $C^{k,Z}$ is locally equivalent to Λ_{k+1} in the sense that any function in one of these spaces agrees on any bounded neighborhood with a function in the other. Similarly for the spaces $C^{k,\alpha}$ and $\Lambda_{k+\alpha}$. Since the Morse-Sard theorem is a local theorem, we are free to use either class. (The Lipschitz classes are more convenient for Theorem 2 since they free us from having to state the integer and noninteger cases separately.)

For any positive integer k, and $0 < \alpha < \beta < 1$, the following inclusions are well-known:

$$\mathbf{C}^{k+1} \subsetneqq \mathbf{C}^{k,1} \subsetneqq C^{k,Z} \subsetneqq \bigcap_{\gamma<1} C^{k,\gamma} \subsetneqq C^{k,\beta} \subsetneqq C^{k,\alpha} \subsetneqq C^k.$$

In terms of the Lipschitz classes, if $0 < s < t$, then

$$\Lambda_t \subsetneqq \Lambda_s.$$

The object of this paper is to prove

Theorem 1 (Zygmund Morse-Sard Theorem). *Let n, m be positive integers with $n > m$. If $f \in C^{n-m,Z}(\mathbf{R}^n, \mathbf{R}^m)$ then the set of critical values of f has Lebesgue measure zero.*

This implies both the classical and Lipschitz Morse-Sard theorems.

In 1953, Dubovickii [3] proved that if $f : \mathbf{R}^n \to \mathbf{R}^m$ is $(n - m + 1)$-times *differentiable*, then the set of its critical values has measure zero. Since differentiable functions need not be Zygmund, this is neither weaker nor stronger than theorem 1. However we note that a Zygmund function can be nowhere differentiable — e.g. the Weierstrass function $f(x) = \sum 2^{-k} \sin(2^k x)$.

Definition. A subset E of $\mathbf{R}^n$ is a *set of rank r for f* if $\text{rank} Df(x) \leq r$ for all $x \in E$.

Theorem 1 is an immediate consequence of

Theorem 2. *Let n, m, and r be nonnegative integers and suppose $n > m > r$. Let E be a set of rank r for $f : \mathbf{R}^n \to \mathbf{R}^m$, and set $s = (n - r)/(m - r)$.*

If $f \in \Lambda_s$, then $f(E)$ has measure zero.

Remarks.

1. Theorem 1 follows from Theorem 2 by setting $r = m - 1$.
2. Bates [1] proved theorem 2 in the case when s is not an integer, making use of a slightly weaker version of the noninteger case in [7].
3. The simple form of theorem 2 is further evidence that the Lipschitz spaces Λ_s are more natural than the smoothness classes C^s for geometric analysis.

The idea of approach to Theorem 2 in the case when s is an integer will be to follow the path of Morse and Sard within the Zygmund class. The key geometric lemma, stated below, permits us to obtain the same estimates as in the Lipschitz case. Then applying an idea of Bates to promote a "big oh" estimate to a "little oh" estimate will complete the proof.

Geometric Lemma. *If $f : \mathbf{R}^n \to \mathbf{R}^m$ is Zygmund with Zygmund constant $C > 0$, and $f(a) = f(b) = 0$, then for all points x belonging to the line segment joining a to b,*

$$|f(x)| \leq (C/2)|b - a|.$$

In Section 1 we present basic facts about the class $C^{k,Z}$ needed in the sequel (relegating the proof of a sharp Composition Theorem to Section 5). Section 2 contains the fundamental Morse Criticality Theorem for the spaces Λ_s. This is used in Section 3 to prove Theorem 2 in the special case of rank $r = 0$. Then in Section 4 we use induction and Fubini's Theorem to arrive at the general case.

Acknowledgement. The author thanks S. Bates, P. Jones, S. Krantz, and D. Sullivan for helpful comments.

1. Basic properties of the Zygmund class

This section contains proofs of basic facts about the class $C^{k,Z}$ that we will need in proving a Zygmund version of the Vanishing Lemma (see Section 2).

We will frequently use the fact that every Zygmund function f has at least a $|t| \log(1/|t|)$ modulus of continuity (e.g. [5]): for every compact set K, there is a constant $C > 0$ such that for all $x \in K$, $|f(x + t) - f(x)| \leq C|t| \log(1/|t|)$ for all t such that $0 < |t| < 1$. Consequently a Zygmund function is always α-Hölder for all $\alpha < 1$.

Product Lemma. *If $f, g : \mathbf{R}^n \to \mathbf{R}$ are of class $C^{k,Z}$, $k \geq 0$, then so is the product fg.*

Proof. First let $k = 0$. For $t \in \mathbf{R}^n$ and $i = 1, 2$, define the first and second difference operators

$$\Delta_t^i : C^0(\mathbf{R}^n, \mathbf{R}) \to C^0(\mathbf{R}^n, \mathbf{R})$$

by $\Delta_t^1 f(x) = f(x+t) - f(x-t)$, and $\Delta_t^2 f(x) \equiv \Delta_t^1 \Delta_t^1 f(x) = f(x+2t) + f(x-2t) - 2f(x)$. Note that f is Zygmund if and only if $\Delta_t^2 f = O(|t|)$ as $|t| \to 0$.

Fix t and let $T_\pm(x) = x \pm t$.

One can check that

$$\Delta_t^1(fg) = (\Delta_t^1 f)(g \circ T_-) + (\Delta_t^1 g)(f \circ T_+)$$

and

$$\Delta_t^2(fg) = f(\Delta_t^2 g) + g(\Delta_t^2 f) + [(\Delta_t^1 f)(\Delta_t^1 g)] \circ T_+ + [(\Delta_t^1 f)(\Delta_t^1 g)] \circ T_-.$$

Now if f and g are Zygmund, the first two terms are $O(|t|)$, and the third term is also $O(|t|)$ because it is a product of two factors, each of which is $O(|t|^\alpha)$ for $\alpha > 1/2$.

This completes the proof for $k = 0$. The statement follows for $k > 0$ by induction from the product rule for differentiation. $\square$

The classes $C^{0,Z}$ and Λ_1 are unfortunately not closed under composition. In fact, even the composition of a Zygmund function with a C^1 function need not be Zygmund (see Section 5). However, if C^1 is strengthened to $C^{1,\alpha}$, we obtain the following statement, sufficient for the purposes of this paper:

Composition Lemma. *For every $m, n, p \in Z^+$ and $\alpha \in (0,1)$,*
(a) $C^{0,Z}(\mathbf{R}^n, \mathbf{R}^m) \circ C^{1,\alpha}(\mathbf{R}^p, \mathbf{R}^n) \subset C^{0,Z}(\mathbf{R}^p, \mathbf{R}^m)$, and
(b) for every $k \in Z^+$, $C^{k,Z}(\mathbf{R}^n, \mathbf{R}^m) \circ C^{k,Z}(\mathbf{R}^p, \mathbf{R}^n) \subset C^{k,Z}(\mathbf{R}^p, \mathbf{R}^m)$.

Proof.
(a) This is an immediate corollary of the sharper Composition Theorem of Section 5.
(b) This follows by induction using the chain rule, part (a), and the Product Lemma above.

Dennis Sullivan pointed out to the author that if f is Zygmund and g is C^1, then $f \circ g$ is Zygmund provided only that Dg satisfies a $1/\log(1/t)$ modulus of continuity, and this is sharp in the sense that no weaker modulus of continuity suffices. Discussion of these technicalities and the statement and proof of a sharp Composition Theorem are relegated to Section 5.

Zygmund Inverse Function Theorem. *If $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R}^n), k \geq 1, x \in \mathbf{R}^n$, and Df_x is a linear isomorphism, then f is invertible in a neighborhood of x and f^{-1} is of class $C^{k,Z}$.*

Proof. The standard proof shows that f^{-1} exists near $f(x)$, is C^k, and satisfies

$$D(f^{-1}) = \mathrm{Inv} \circ Df \circ f^{-1},$$

where Inv denotes the inverse operator on GL(n). In fact, for any $\alpha \in (0,1)$, f^{-1} is $C^{k+\alpha}$ by the $C^{k+\alpha}$ Inverse Function Theorem [7]. Therefore, since Inv is C^{∞}, Df is $C^{k-1,Z}$, and f^{-1} is $C^{k+\alpha}$, the Composition Lemma implies that $D(f^{-1})$ is $C^{k-1,Z}$. $\square$

Combining this with the $C^{k+\alpha}$ Inverse Function Theorem [7], we obtain the general

Λ_s **Inverse Function Theorem.** *If $f \in \Lambda_s(\mathbf{R}^n, \mathbf{R}^n), s > 1, x \in \mathbf{R}^n$, and Df_x is a linear isomorphism, then f is invertible in a neighborhood of x and f^{-1} is of class Λ_s.*

Many standard arguments now go through easily in the Zygmund class, e.g. the

Zygmund Preimage Theorem. *If $k \geq 1$, $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R}), x \in \mathbf{R}^n, f(x) = 0$, and $Df(x) \neq 0$, then there is a neighborhood N of x in $\mathbf{R}^n$ and a $C^{k,Z}$ $(n-1)$-submanifold $S \subset \mathbf{R}^n$ such that*

$$f^{-1}(0) \cap N \subset S.$$

Proof. Use the standard argument (e.g. [4]) with the Zygmund Inverse Function Theorem.

To finish this section, we restate and prove the

Geometric Lemma. *If $f : \mathbf{R}^n \to \mathbf{R}^m$ is Zygmund with Zygmund constant $C > 0$, $a, b \in \mathbf{R}^n$, and $f(a) = f(b) = 0$, then for all points x belonging to the line segment $[a, b]$ joining a to b,*

$$|f(x)| \leq (C/2)|b - a|.$$

Proof. We are supposing that for all $x, h \in \mathbf{R}^n$,

$$|f(x+h) + f(x-h) - 2f(x)| \leq C|h|.$$

If I is a line segment with endpoints y and z and midpoint x, this implies that

$$|f(x)| \leq C|I|/4 + (|f(y)| + |f(z)|)/2. \tag{1}$$

Let $L = |b - a|$. For $k = 0, 1, 2, \ldots$, consider the collection $\mathcal{C}_k$ of closed dyadic subintervals of $[a, b]$, i.e. intervals with endpoints $a + (p/2^k)(b - a)$ and $a + ((p + 1)/2^k)(b - a)$, where p is an integer between 0 and $2^k - 1$.

Define $B_0 = \{a, b\}$, and for $k = 1, 2, 3, \ldots$, define B_k to be the finite set of points x in $[a, b]$ such that x is an endpoint of an interval of $\mathcal{C}_k$ but is an endpoint of no interval of $\mathcal{C}_j$ for any $j < k$. (B_k is the set of endpoints that "arise at stage k".)

Claim: For all $x \in B_k$, $|f(x)| \leq CL(1 - 2^{-k})/2$.

We prove the claim below, but first note that this implies $|f(x)| \leq CL/2$ for all x in the dense set $\cup_k B_k$. Since f is continuous, this yields the conclusion of the Geometric Lemma.

The claim is proved by simple induction. The case $k = 0$ is simply our hypothesis that $f(a) = f(b) = 0$. Let $k > 0$ and suppose that for $j = 0, 1, 2, \ldots, k - 1$,

$$|f(y)| \leq CL(1 - 2^{-j})/2 \text{ for all } y \in B_j.$$

Take $x \in B_k$. Then x is the midpoint of some dyadic interval of length $L2^{-(k-1)}$, with endpoints, say, x' and x''. By (1) above,

$$|f(x)| \leq CL/2^{k+1} + (|f(x')| + |f(x'')|)/2.$$

By induction, since x' and x'' must appear at some stage before the kth, $|f(x')|$ and $|f(x'')|$ are no more than $CL(1-2^{-(k-1)})/2$. Therefore

$$\begin{aligned} |f(x)| &\leq CL/2^{k+1} + CL(1-2^{-(k-1)})/2 \\ &= CL(1-2^{-k})/2. \end{aligned}$$

□

2. A Zygmund version of the Morse Criticality Lemma

In this section we take the fundamental step in the proof of Theorem 2 by proving the

Morse Criticality Theorem. *Let $s \geq 1$ be a real number, $n \in Z^+$, $A \subset \mathbf{R}^n$. Then there is a countable collection $\{A_i\}$ of bounded subsets of A such that*

(i) $A = \cup A_i$, and

(ii) for every $f \in C^1 \cap \Lambda_s(\mathbf{R}^n, \mathbf{R})$ critical on A, and every i, there is a constant $c_i > 0$ such that

$$\textit{for every } x, y \in A_i, \quad |f(x) - f(y)| \leq c_i |x-y|^s.$$

In case $s = 1$ or $s \notin Z^+$, this theorem appears in [7]. The only remaining case is the

Zygmund Morse Criticality Lemma. *Let n, k be positive integers and A a subset of $\mathbf{R}^n$. Then there is a countable collection $\{A_i\}$ of bounded subsets of A such that*

(i) $A = \cup A_i$, and

(ii) for every $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R})$ critical on A and every i, there is a constant $c_i > 0$ such that

$$\textit{for every } x, y \in A_i, \quad |f(x) - f(y)| \leq c_i |x-y|^{k+1}.$$

This will be a direct consequence of the following lemma.

Definitions. We will say that the pair (B, ϕ) is a C^1 *parametrized disk in* $\mathbf{R}^n$ if for some integer p, $1 \leq p \leq n$, ϕ is a C^1 embedding of $\mathbf{R}^p$ into $\mathbf{R}^n$ and $\phi(\Delta) = B$, where Δ is the closed unit ball in $\mathbf{R}^p$.

Given such a parametrized disk (B, ϕ), and two points $x, y \in B$, the *parametric segment in B from x to y* is defined to be the image under ϕ of the line segment joining $\phi^{-1}(x)$ to $\phi^{-1}(y)$.

Zygmund Morse Vanishing Lemma. *Let n be a positive integer, k a non-negative integer, and A a subset of $\mathbf{R}^n$. Then there exists a countable set A_0, a countable collection $\{A_i\}_{i=1}^{\infty}$ of bounded subsets of $\mathbf{R}^n$, and a countable collection $\{(B_i, \phi_i)\}_{i=1}^{\infty}$ of C^1 parametrized disks in $\mathbf{R}^n$ such that*

(i) $A = \cup_{i=0}^{\infty} A_i$,

(ii) for each $i \in Z^+$, $A_i \subset B_i$, and

(iii) For every $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R})$ vanishing on A, and every $i \in Z^+$, there is a positive constant c_i such that

$$|f(z)| \le c_i |x - z|^k |x - y|$$

whenever $x, y \in A_i$ and z lies on the parametric segment in B_i from x to y.

We will call such a collection $\{A_i\}$ for a set A a *Morse decomposition of A*.

Proof.

The proof is by double induction on n and k. Let $\langle n, k\rangle$ stand for the statement of the theorem for $\mathbf{R}^n$ and $C^{k,Z}$. We will prove (a) $\langle n, 0\rangle$ for all n, (b) $\langle 1, k\rangle$ for all k, and (c) $\langle n-1, k\rangle$ and $\langle n, k-1\rangle$ imply $\langle n, k\rangle$.

(a) *Proof of $\langle n, 0\rangle$ for all n.*

If $A \subset \mathbf{R}^n$, let $A_0 = \emptyset$, B_i be the closed ball in $\mathbf{R}^n$ with center 0 and radius i, ϕ_i be the linear expansion $x \mapsto ix$ on $\mathbf{R}^n$, and $A_i = A \cap B_i$.

Given $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R})$ vanishing on A, for each i take c_i to be the Zygmund constant of f for the compact set B_i. The conclusion follows immediately from the Geometric Lemma of Section 1.

(b) *Proof of $\langle 1, k\rangle$ for all k.*

Let A be a subset of $\mathbf{R}$. Denote by A^* the set of condensation points of A; that is, A^* is the set of points x in A such that every neighborhood of x meets A in uncountably many points. It is easy to show that $A_0 \equiv A \setminus A^*$ is countable.

Now for $i > 0$ let $A_i = A^* \cap (-i, i)$, $B_i = [-i, i]$, and ϕ_i be the linear expansion by factor i. Given $f \in C^{k,Z}(\mathbf{R}, \mathbf{R})$, let c_i' be the Zygmund constant of $D^k f$ for the compact set B_i.

Since every point of A_i is a limit point of A_i, and $f \in C^1$ vanishes on A_i, Df must vanish at every point of A_i. Similarly, $D^2 f, D^3 f, \ldots, D^k f$ all vanish on A_i.

By the Geometric Lemma applied to $D^k f$ on B_i, if $x, y \in A_i$ and $x < t < y$, then

$$|D^k f(t)| \le c_i' |x - y|.$$

Hence, for any z between x and y, by integrating $Df(t)$ k times from x to z we obtain,

$$\begin{aligned}
|f(z)| &= \left| \int_x^z \int_x^{t_k} \cdots \int_x^{t_2} D^k f(t_1) dt_1 dt_2 \ldots dt_k \right| \\
&\le \int_x^z \int_x^{t_k} \cdots \int_x^{t_2} c_i' |x - y| dt_1 dt_2 \ldots dt_k \\
&= c_i |x - z|^k |x - y|,
\end{aligned}$$

where $c_i = c_i'/k!$.

(c) Induction step: we assume $\langle n-1, k\rangle$ and $\langle n, k-1\rangle$, and prove $\langle n, k\rangle$.

Define

$$U = \{x \in A: \text{ for every } g \in C^{k,Z}(\mathbf{R}^n, \mathbf{R}) \text{ vanishing on } A, \quad Dg(x) = 0\}, \text{ and}$$
$$V = A \setminus U.$$

Since the union of two Morse decompositions is again a Morse decomposition, it suffices to prove the result for U and V separately.

We first address U. By our $\langle n, k-1 \rangle$ hypothesis, there are subsets $U_0, U_1, U_2, \ldots$ of U and C^1 parametrized disks (B_i, ϕ_i) such that U_0 is countable, $U = \cup U_i$, and $U_i \subset B_i$ for $i \geq 1$.

Moreover for any $h \in C^{k-1,Z}(\mathbf{R}^n, \mathbf{R})$ vanishing on V, there are constants $c_i > 0$ such that

$$|h(t)| \leq c_i |x-t|^{k-1} |x-y|$$

whenever $x, y \in U_i$ and t lies on the parametric segment $[x, y]_{B_i}$ in B_i from x to y.

Now suppose $f \in C^k(\mathbf{R}^n, \mathbf{R})$ vanishes on U. By definition of U, f must be critical on U, so $D_j f \equiv \partial f / \partial x_j$ vanishes on U for $j = 1, \ldots, n$. Since these functions lie in $C^{k-1}(\mathbf{R}^n, \mathbf{R})$, by above we know there exist constants $c_{i,j}$, $i = 1, 2, 3, \ldots; j = 1, 2, \ldots, n$, so that

$$|D_j f(t)| \leq c_{i,j} |x-t|^{k-1} |x-y|,$$

hence

$$||Df(t)|| \leq (\max_j c_{i,j}) |x-t|^{k-1} |x-y|, \tag{2}$$

whenever $x, y \in U_i$ and t lies on the parametric segment $[x, y]_{B_i}$.

Now we need the following lemma, whose proof is straightforward and omitted:

Lemma 1. *If Δ denotes the closed unit ball in $\mathbf{R}^p$ and ϕ is a C^1 embedding of Δ into $\mathbf{R}^n$, then there is a constant κ, depending on ϕ, such that*

$$(1/\kappa)|a-b| \leq |\phi(a) - \phi(b)| \leq \kappa |a-b|$$

for all $a, b \in \Delta$.

Applying Lemma 1 to ϕ_i and B_i, we obtain a constant k_i satisfying the conclusion of Lemma 1, so that for all $x, z \in B_i$ and $t \in [x, z]_{B_i}$, we have

$$\begin{aligned} |x-t| &\leq k_i |\phi_i^{-1}(x) - \phi_i^{-1}(t)| \\ &\leq k_i |\phi_i^{-1}(x) - \phi_i^{-1}(z)| \\ &\leq k_i^2 |x-z|. \end{aligned}$$

Therefore, letting $c_i' = (\max_j c_{i,j}) k_i^{2(k-1)}$, and fixing z on $[x, y]_{B_i}$, (2) yields

$$||Df(t)|| \leq c_i' |x-z|^{k-1} |x-y|$$

for all $t \in [x, z]_{B_i}$.

Now integrating Df along $[x, y]_{B_i}$ from x to z, we obtain

$$|f(z)| \leq c_i' |x-z|^{k-1} |x-y| L(x, z),$$

where $L(x,z)$ denotes the length of the parametric segment $[x,z]_{B_i}$.

It follows from Lemma 1 that there is a constant $b_i > 0$ such that $L(x,z) \le b_i|x-z|$ for all $x, z \in B_i$. Therefore we may let $c_i = c_i' d_i$ and deduce that

$$|f(z)| \le c_i |x-z|^k |x-y|$$

whenever $x, y \in U_i$ and z lies on the parametric segment from x to y in B_i. This is our desired conclusion for U.

Turning now to V, note first that it suffices to show that V satisfies the conclusion locally, i.e. in a neigborhood of each of its points. This is because a countable union of Morse decompositions is again a Morse decomposition.

So given $x \in V$, we show that, for some neighborhood N of x in $\mathbf{R}^n$, $V \cap N$ has an appropriate Morse decomposition.

Since $x \in V$, some $g \in C^{k,Z}(\mathbf{R}^n, \mathbf{R})$ vanishing on A is not critical at x. By the Zygmund Preimage Theorem, there exists a neighborhood N of x such that $A \cap N$, and hence $V \cap N$, is contained in a $C^{k,Z}$ $(n-1)$-submanifold S (namely, S is the local zero set for g).

We can choose S and an embedding $\psi \in C^{k,Z}(\mathbf{R}^{n-1}, \mathbf{R}^n)$ so that

$$V \cap N \subset S = \psi(\Delta),$$

where Δ denotes the closed unit ball in $\mathbf{R}^{n-1}$.

Writing $W = \psi^{-1}(V \cap N)$, we now have, by our $\langle n-1, k\rangle$ hypothesis:

(i) a collection $\{W_i\}_{i=0}^{\infty}$ of subsets of W, with W_0 countable and $W = \cup W_i$,

(ii) a collection $\{(D_i, \eta_i)\}$ of C^1 parametrized disks in $\mathbf{R}^{n-1}$ so that $W_i \subset D_i$ for all $i \ge 1$, and

(iii) for every $h \in C^{k,Z}(\mathbf{R}^{n-1}, \mathbf{R})$ vanishing on W, and every $i \in Z^+$, a positive constant d_i such that

$$|h(z)| \le d_i |x-z|^k |x-y|$$

whenever $x, y \in W_i$ and z lies on the parametric segment $[x,y]_{D_i}$ in D_i joining x to y.

Now define, for all i,

$$V_i = \psi(W_i),\ B_i = \psi(D_i),\ \text{and}\ \phi_i = \psi \circ \eta_i.$$

Then the following facts are immediate:

(i') V_0 is countable and $\cup V_i = V \cap N$, and

(ii') each pair (B_i, ϕ_i) is a C^1 parametrized disk in $\mathbf{R}^n$ with $V_i \subset B_i$ for all $i \ge 1$.

Moreover,

(iii') if $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R})$ vanishes on $V \cap N$, then $f \circ \psi \in C^{k,Z}(\mathbf{R}^{n-1}, \mathbf{R})$ vanishes on W.

If we pick an arbitrary $x', y' \in V_i$ and z' on the parametric segment $[x', y']_{B_i}$, then there are points x, y, z in D_i such that $\psi(x) = x', \psi(y) = y'$, $\psi(z) = z'$, $x, y \in W_i$, and $z \in [x,y]_{D_i}$.

Hence, by (iii) above, there exist constants d_i, depending on f but independent of x, y, z, such that

$$|f \circ \psi(z)| \le d_i |x-z|^k |x-y|.$$

Applying Lemma 1 to the embedding ψ, we obtain a constant $k_i > 0$ such that

$$|a - b| \leq k_i |\psi(a) - \psi(b)|$$

for all $a, b \in \Delta$. In particular, $|x - z| \leq k_i |x' - z'|$ and $|x - y| \leq k_i |x' - y'|$, so that

$$|f(z')| \leq d_i k_i^{k+1} |x' - z'|^k |x' - y'|.$$

This completes the inductive step, and the proof. $\square$

Proof of Zygmund Morse Criticality Lemma.

Let $\{A'_i\}$ be the Morse decomposition given by the Morse Vanishing Lemma for the pair $\langle n, k-1 \rangle$. Let $\{(B_i, \phi_i)\}$ be the corresponding C^1 parametrized disks. We check (ii).

Given $f \in C^{k,Z}(\mathbf{R}^n, \mathbf{R})$ critical on A, this means that for each $j = 1, \ldots, n$, $D_j f$ lies in $C^{k-1,Z}$ and vanishes on A.

The Morse Vanishing Lemma now guarantees us constants $c_{i,j} > 0$ such that

$$|D_j f(z)| \leq c_{i,j} |x - z|^{k-1} |x - y|$$

whenever $x, y \in A'_i$ and $z \in [x, y]_{B_i}$.

By Lemma 1, as in the preceding argument, there is a constant $k_i > 0$ such that $|x - z| \leq k_i^2 |x - y|$ so long as $x, y \in A'_i$ and $z \in [x, y]_{B_i}$. Therefore $|Df(z)|$ is controlled by a constant times $|x - y|^k$. Integrating Df along $[x, y]_{B_i}$ from x to z gives, as in the proof of the Vanishing Lemma, a constant $c_i > 0$, depending only on f and i, such that

$$|f(x) - f(y)| \leq c_i |x - y|^{k+1}$$

whenever $x, y \in A'_i$.

We now have a collection $\{A'_i\}_{i=1}^{\infty}$ satisfying the desired conclusions, except that the union $\cup A'_i$ omits the countable subset A'_0 of A. Let $\{A_i\} = \{A'_i\} \cup \{\{x\} : x \in A'_0\}$. This countable collection now suffices, since when A_i is a singleton, the condition (ii) is trivially true. $\square$

3. The Rank Zero Lemmas

The following lemmas will establish the rank zero case of Theorem 2. In section 4 we repeat the argument in [7] and [10] deducing the general case from this one.

First Rank Zero Lemma. *Let $n > m$, $f \in \Lambda_{n/m}(\mathbf{R}^n, \mathbf{R}^m)$, and $A \subset \mathbf{R}^n$ be a set of rank zero for f. If A has measure zero, then $f(A)$ has measure zero.*

Proof. By the Morse Criticality Theorem, there is a Morse decomposition $A = \cup A_i$ such that every component of f, and therefore f itself, satisfies

$$|f(x) - f(y)| \leq c_i |x - y|^{n/m} \tag{2}$$

for some $c_i > 0$ and every $x, y \in A_i$, $i > 0$.

Fix j and write $B = A_j$ and $c = c_j$; we show that $f(B)$ has measure zero. Let $\epsilon > 0$ be given. Since B has measure zero, it is contained in the union $\cup B_i$ of balls in $\mathbf{R}^n$ chosen such that $\sum |B_i|^n < \epsilon / c^m$, where $|B_i|$ denotes the diameter of B_i.

By (2), for each i,

$$|f(B \cap B_i)| \leq c|B \cap B_i|^{n/m}.$$

Letting l_m denote m-dimensional Lebesgue measure, we therefore have

$$\begin{aligned} l_m(f(B)) &\leq \sum l_m(f(B \cap B_i)) \leq \sum |f(B \cap B_i)|^m \\ &\leq \sum (c|B \cap B_i|^{n/m})^m \\ &\leq c^m \sum |B_i|^n < \epsilon. \end{aligned}$$

Hence $l_m(f(B)) = 0$. $\square$

In the next lemma we invoke an argument of Bates [1] to remove the requirement in the First Rank Zero Lemma that A have measure zero.

Second Rank Zero Lemma. *Let $n > m$, $f \in \Lambda_{n/m}(\mathbf{R}^n, \mathbf{R}^m)$, and $A \subset \mathbf{R}^n$ be a set of rank zero for f.*

Then $f(A)$ has measure zero.

Proof. Let $\{A_i\}$ be a Morse decomposition for A as provided by the Morse Criticality Lemma. (Without loss of generality, we assume that $l_n(A_i) < \infty$ for each i.) Fix $j \geq 1$ and write $B = A_j$; our job is to show that $l_m(f(B)) = 0$.

By the Morse Criticality Lemma applied to each component of f, there is a constant $c > 0$ such that

$$|f(x) - f(y)| \leq c|x - y|^{n/m} \tag{3}$$

for all $x, y \in B$.

We may assume that B is measurable, since replacing B by its closure preserves condition (3). Hence, by the Lebesgue Density Theorem, almost every point of B is a density point. That is, we can write $B = D \cup E$, where $l_n(E) = 0$ and every point of D is a density point of D. The First Rank Zero Lemma shows that $l_m(f(E)) = 0$; it remains to show that $l_m(f(D)) = 0$.

The following lemma is convenient:

Lemma 2. *If $n \geq 1, P \geq 2$ are integers, $D \subset \mathbf{R}^n$ is measurable, and Q is a cube in $\mathbf{R}^n$ satisfying*

$$\frac{l_n(D \cap Q)}{l_n(Q)} \geq 1 - P^{-n}, \tag{4}$$

then for every $x, y \in D \cap Q$, there is a sequence $x_0, \ldots, x_P$ of points in $D \cap Q$ such that $x_0 = x$, $x_P = y$, and

$$|x_i - x_{i+1}| < 2|Q|/P$$

for $i = 0, \ldots, P - 1$.

Lemma 2 is proved below, but first we finish the proof of the Second Rank Zero Lemma. First choose an arbitrary $P \geq 2$. Since every point of D is a density point, there is a positive function $\delta : D \to \mathbf{R}$ so that any cube $Q(x, \delta)$ with center x and side length $\delta < \delta(x)$ satisfies (4).

Let Q be any such cube. Then for every $x, y \in D \cap Q$,

$$\begin{aligned}|f(x) - f(y)| &\leq |f(x_0) - f(x_1)| + \cdots + |f(x_{P-1}) - f(x_P)| \\ &\leq c|x_0 - x_1|^{n/m} + \cdots + c|x_{P-1} - x_P|^{n/m} \text{ (by (3))} \\ &\leq cP(2|Q|/P)^{n/m} \\ &= c2^{n/m}P^{1-(n/m)}|Q|^{n/m}.\end{aligned}$$

This means that $|f(D \cap Q)| \leq c2^{n/m}P^{1-(n/m)}|Q|^{n/m}$, so

$$\begin{aligned}l_m(f(D \cap Q)) \leq |f(D \cap Q)|^m &\leq c^m 2^n P^{m-n}|Q|^n \\ &\leq c^m 2^{2n} P^{m-n} l_n(Q).\end{aligned}$$

Since $\{Q(x, \delta) : x \in D, \delta < \delta(x)\}$ is a Vitali family for D, there is a countable subcollection $\{Q_i\}$ such that $l_n(D \setminus \cup Q_i) = 0$ and $\sum l_n(Q_i) < 2l_n(D)$. ($l_n(D) < \infty$ since D is bounded.) By the First Rank Zero Lemma, $l_m(f(D \setminus \cup Q_i)) = 0$.

Hence

$$\begin{aligned}l_m(f(D)) &\leq \sum l_m(f(Q_i)) \\ &\leq \sum c^m 2^{2n} P^{m-n} l_n(Q_i) \\ &< 2^{2n+1} c^m P^{m-n} l_n(D).\end{aligned}$$

Since P is arbitrary, $l_m(f(D))$ must vanish, and this completes the proof. □

Proof of Lemma 2. Assume (4) and $x, y \in D \cap Q$. Clearly the line segment L joining x to y can be covered by at most P subcubes of Q of diameter $|Q|/P$. By (4), each of these subcubes must contain a point of $D \cap Q$, so choosing a point in each yields the desired sequence. □

4. Proof of Theorem 2

For nonnegative integers $n > m > r$, let $s = (n - r)/(m - r)$. We assume $f \in \Lambda_s(\mathbf{R}^n, \mathbf{R}^m)$ and $E \subset \mathbf{R}^n$ is a set of rank r for f.

For $i = 0, 1, 2, \ldots, r$, define $R_i = \{x \in E : \text{rank} Df(x) = i\}$. We need to show that $l_m(f(R_i)) = 0$ for each i.

The case $i = 0$ is covered by the Second Rank Zero Theorem of the previous section, so we now fix $i \geq 1$. It will suffice to find, for every $p \in R_i$, a neighborhood U of p such that $l_m(f(U \cap R_i)) = 0$.

By means of a standard argument (e.g. as in [4]) using the Λ_s Inverse Function Theorem, we can find coordinates in some neighborhood U of p so that

$$f(x_1, \ldots, x_n) = (x_1, \ldots, x_i, g(x_1, \ldots, x_n)),$$

where $g \in \Lambda_s(\mathbf{R}^n, \mathbf{R}^{m-i})$.

In these coordinates,

$$Df(x) = \begin{pmatrix} Id_i & 0 \\ * & D(g[x_1, \ldots, x_i]) \end{pmatrix},$$

where $x = (x_1, \ldots, x_n)$, Id_i is the $i \times i$ identity matrix, and $g[x_1, \ldots, x_i]$ denotes the function $\mathbf{R}^{n-i} \to \mathbf{R}^{m-i}$ defined by

$$g[x_1, \ldots, x_i] : (x_{i+1}, \ldots, x_n) \mapsto g(x_1, \ldots, x_n).$$

By definition of R_i, if $x \in U \cap R_i$, then $\mathrm{rank} D(g[x_1, \ldots, x_i]) = 0$.

Define the "cross-section" of a set $A \subset \mathbf{R}^n$ at $(x_1, \ldots, x_i) \in \mathbf{R}^i$ by

$$A[x_1, \ldots, x_i] = \{(x_{i+1}, \ldots, x_n) \in \mathbf{R}^{n-i} : (x_1, \ldots, x_n) \in A\}.$$

In these terms, it is easy to check that, for $x \in U \cap R_i$, $g[x_1, \ldots, x_i]$ maps the rank 0 set $(U \cap R_i)[x_1, \ldots, x_i] \subset \mathbf{R}^{m-i}$ onto the set $(f(U \cap R_i))[x_1, \ldots, x_i] \subset \mathbf{R}^{m-i}$.

Now, because $(n-i)/(m-i) \leq (n-r)/(m-r) = s$, we have

$$g[x_1, \ldots, x_i] \in \Lambda_s(\mathbf{R}^{n-i}, \mathbf{R}^{m-i}) \subset \Lambda_{(n-i)/(m-i)}(\mathbf{R}^{n-i}, \mathbf{R}^{m-i}).$$

Therefore, by the Second Rank Zero Lemma applied to $g[x_1, \ldots, x_i]$ on $(U \cap R_i)[x_1, \ldots, x_i]$, we obtain

$$l_{m-i}((f(U \cap R_i))[x_1, \ldots, x_i]) = 0.$$

Since this holds for every $x \in U \cap R_i$, we may apply Fubini's Theorem for $\mathbf{R}^m = \mathbf{R}^i \times \mathbf{R}^{m-i}$ to deduce

$$l_m(f(U \cap R_i)) = 0.$$

□

5. Moduli of Continuity and a Composition Theorem

In this section we discuss, as promised in Section 1, the question of how smooth a function g must be so that $f \circ g$ is guaranteed to be Zygmund if f is Zygmund. For this we need to discuss moduli of continuity.

Definition. If $(X, d), (Y, d')$ are metric spaces and $f : X \to Y$ is continuous, the *modulus of continuity* m_f *of* f is defined to be the function

$$m_f(t) = \sup\{d'(f(x), f(y)) : d(x, y) \leq t\}$$

for $t \geq 0$. (This function might in general be infinite for some or all $t > 0$.)

A function $\alpha : [0, \infty) \to [0, \infty]$ is called a *modulus of continuity* if α is monotone, continuous, and $\alpha(0) = 0$.

These definitions are consistent because of the fact (whose proof we omit) that for every modulus of continuity α, there is a (possibly infinite-dimensional) metric space (X, d) and a continuous function $f : X \to \mathbf{R}$ such that $m_f = \alpha$.

We now restrict attention to functions defined on Euclidean spaces. Since we are interested in local smoothness properties, we are free to confine our attention to continuous functions $f : \mathbf{R}^n \to \mathbf{R}^m$ with compact support (i.e. locally constant outside a compact set). Such functions always have a finite modulus of continuity.

Definition. A function f is *majorized* by a modulus of continuity α if $m_f \leq \alpha$. Equivalently,

$$|f(x) - f(y)| \leq \alpha(|x-y|)$$

for all x, y.

Example. If f is majorized by $\alpha(t) = ct$, then f is Lipschitz; if by $\alpha(t) = ct^s$, then f is s-Hölder.

Only certain functions α can arise as moduli of continuity for functions $f : \mathbf{R}^n \to \mathbf{R}^m$ of compact support —these are the *subadditive* functions, i.e. those satisfying

$$\alpha(s+t) \leq \alpha(s) + \alpha(t)$$

for all $s, t \in [0, \infty)$.

It is easy to check that for any such f, m_f is subadditive and has compact support; conversely, if α is a subadditive modulus of continuity with compact support, then $\alpha = m_f$ for some f —namely, $f(t) = \alpha(t)$ for $t \geq 0$, $f(t) = 0$ otherwise.

Definition. For $k = 0, 1, 2, \ldots$ and α a modulus of continuity, we say that

$$f \in C^{k,\alpha}$$

if $f \in C^k$ and $D^k f$ is locally majorized by a multiple of α. If f also has compact support, we write

$$f \in C_c^{k,\alpha}.$$

Suppose α and β are moduli of continuity, and for some $C > 1$ and all $t > 0$,

$$1/C \leq \frac{\alpha(t)}{\beta(t)} \leq C.$$

We then think of α and β as equivalent because $C^{k,\alpha} = C^{k,\beta}$. This defines an equivalence relation $\sim$ on the set of all moduli of continuity. There is a natural partial order $\preccurlyeq$:

$$\alpha \preccurlyeq \beta \text{ if } \alpha \sim \beta \text{ or } \alpha \prec \beta,$$

where

$$\alpha \prec \beta \text{ means } \alpha(t)/\beta(t) \to \infty \text{ as } t \to 0^+.$$

Clearly, if $\alpha \preccurlyeq \beta$, then $C_c^{k,\beta} \subset C_c^{k,\alpha}$.

Now we can state the

Composition Theorem. *Let*

$$\lambda(t) = \begin{cases} \frac{1}{\log(1/t)} & \textit{for } 0 < t \leq e^{-2} \\ 1/2 & \textit{for } t \geq e^{-2} \end{cases}$$

and $\lambda(0) = 0$.

Then λ is a subadditive modulus of continuity. Moreover,

(a) If $f \in C^{0,Z}(\mathbf{R}^n, \mathbf{R}^m)$ and $g \in C^{1,\lambda}(\mathbf{R}^p, \mathbf{R}^n)$, then $f \circ g \in C^{0,Z}$, and

(b) There exists $f \in C^Z(\mathbf{R}, \mathbf{R})$ such that for every subadditive modulus of continuity $\alpha \prec \lambda$, there is a function $g \in C_c^{1,\alpha}(\mathbf{R}, \mathbf{R})$ such that $f \circ g \notin C^Z$.

Proof. One easily checks that λ is continuous, monotone, and concave, hence subadditive.

Proof of (a). Given $f \in C^Z, g \in C^{1+\lambda}$, we show that the composition $f \circ g$ is Zygmund by a straightforward estimate of the second difference.

Fix a convex compact set K in $\mathbf{R}^p$. Denote by A the Zygmund constant of f on $g(K)$, and choose B so that $|Dg(x) - Dg(y)| \leq B\lambda(|x-y|)$ for $x, y \in K$. Recall that f must satisfy a $t\log(1/t)$ modulus of continuity. That is, if we let $\beta(t) = t\log(1/t)$ for $0 < t < 1/2$, and $\beta(t) = (\log 2)/2$ otherwise, then there exists $L > 0$ such that $|f(x) - f(y)| \leq L\beta(|x-y|)$ for all $x, y \in K$.

Applying Taylor's theorem to the components g_i of g, we find that

$$g_i(x+t) = g_i(x) + Dg_i(\xi_i)t,$$

where ξ_i is some point on the line segment $[x, x+t]$, $x, x+t \in K$.

Hence

$$\begin{aligned}|g(x+t) - g(x) - Dg(x)t| &\leq \sum_i |g_i(x+t) - g_i(x) - Dg_i(x)t| \\ &\leq \sum |Dg_i(\xi_i)t - Dg_i(x)t| \\ &\leq nB\lambda(|t|)|t|.\end{aligned}$$

Now we estimate:

$$\begin{aligned}&|f \circ g(x+t) + f \circ g(x-t) - 2f \circ g(x)| \\ &\leq |f(g(x) + Dg(x)t) + f(g(x) - Dg(x)t) - 2f(g(x))| \\ &\qquad + |f(g(x) + Dg(x)t) - f(g(x+t))| + |f(g(x) - Dg(x)t) - f(g(x-t))| \\ &\leq A\|Dg(x)\||t| + L\beta(|g(x+t) - g(x) - Dg(x)t|) \\ &\qquad + L\beta(|g(x-t) - g(x) - Dg(x)(-t)|) \\ &\leq A\|Dg(x)\||t| + 2L\beta(nB\lambda(|t|)|t|).\end{aligned}$$

The first term on the right is $O(|t|)$ as $|t| \to 0$ since $\|Dg\|$ is bounded on K. That the second term is also $O(|t|)$, with constants depending only on K, A, B and L, is easily verified using the definitions of β and λ, and the fact that $(\log(C/(-\log|t|)))/\log|t|$ tends to zero as $|t| \to 0$.

Proof of (b). Let $f(x) = \sum_{k=0}^{\infty} 2^{-k}\sin(2^k x)$.

It is known that f is Zygmund (e.g. [5]), and we use the fact that there exists $C > 0$ such that

$$f(x) \geq Cx\log(1/x)$$

for $0 < x < e^{-2}$.

Given the subadditive modulus of continuity α, extend α to all of $\mathbf{R}$ by letting it have value zero on $\mathbf{R}^-$. Construct $\hat{\alpha}$ by reflection in the line $x = 1$:

$$\hat{\alpha}(t) = \begin{cases} \alpha(t) & \text{for } t \leq 1 \\ \alpha(1-t) & \text{for } t > 1. \end{cases}$$

Now define $g(x) = \int_0^x \hat{\alpha}(t)\,dt$. Evidently g has compact support since $\hat{\alpha}$ is zero outside of [0,2]. To show that $g \in C^{1,\alpha}$, we need only show that $\hat{\alpha}$ is majorized by α.

Case 1: $0 \le x < y \le 1$. Then

$$|\hat{\alpha}(x) - \hat{\alpha}(y)| = |\alpha(x) - \alpha(y)| = |\alpha(x) - \alpha(x + (y - x))| \le \alpha(y - x)$$

by subadditivity.

Case 2: $1 \le x < y$. Similar.

Case 3: $x < 1 < y$. Suppose x is closer to 1 than y (otherwise argue similarly). Then

$$|\hat{\alpha}(x) - \hat{\alpha}(y)| \le |\hat{\alpha}(1) - \hat{\alpha}(y)| \le \alpha(|1 - y|)$$

as before, and by monotonicity this is at most $\alpha(|x - y|)$.

It now remains to establish that $f \circ g$ is not Zygmund. Choose a positive integer N. Since $\alpha \prec \lambda$, there is $\epsilon = \epsilon(N) < e^{-2}$ such that $0 < x < \epsilon$ implies $\alpha(t) \ge N\lambda(t)$.

For such x,

$$\begin{aligned}
|f(g(x)) + f(g(-x)) - 2f(g(0))| &= |f(g(x))| \\
&\ge Cg(x)\log(1/g(x)) \\
&\ge C(\int_0^x N\lambda(t)dt)(-\log\int_0^x N\lambda(t)dt)
\end{aligned}$$

since the function $t\log(1/t)$ is monotone on $(0, e^{-2})$,

$$= CN(\int_0^x \lambda(t)dt)[-\log N - \log\int_0^x \lambda(t)dt].$$

If we choose $\epsilon' \le \epsilon$ so small that $x < \epsilon'$ implies

$$\int_0^x \lambda(t)dt < 1/N^2,$$

then

$$-\log\int_0^x \lambda(t)dt > 2\log N,$$

and so

$$f(g(x)) \ge (CN/2)(\int_0^x \lambda(t)dt)(-\log\int_0^x \lambda(t)dt).$$

Now denote by $\bar{\lambda}(x)$ the average value $(1/x)\int_0^x \lambda(t)dt$ of λ on the interval $[0, x]$.

We now obtain

$$\begin{aligned}
f(g(x))/x &\ge (CN/2)\{\bar{\lambda}(x)[-\log\bar{\lambda}(x) + \log(1/x)]\} \\
&= (CN/2)\{\bar{\lambda}(x)\log(1/\bar{\lambda}(x)) + \bar{\lambda}(x)\log(1/x)\}.
\end{aligned}$$

The first term on the right tends to zero as x tends to zero. The second term is

$$\frac{\int_0^x \lambda(t)dt}{x\lambda(x)}.$$

One easily checks, using L'Hopital's rule and the definition of λ, that this tends to one as x tends to zero.

Hence, for x sufficiently small,

$$\Delta^2_{(x/2)} f(0) = f(g(x)) \geq (CN/4)x.$$

Since N was chosen arbitrarily, this means that $f(g(x))/x \to +\infty$ as $x \to 0^+$, so $f \circ g$ is not Zygmund.

References

1. S. Bates, *Toward a precise smoothness hypothesis in Sard's Theorem*, Proc. Amer. Math. Soc. **117** (1993), 279–283.
2. ———, *On the image size of singular maps, II*, Duke Math. J. **68** (1992), 463–476.
3. A. Ya. Dubovickii, *On differentiable mappings of an n-dimensional cube into a k-dimensional cube*, Mat. Sbornik (N.S.) **32** (1953), 443–464. (Russian)
4. V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, 1974.
5. S. Krantz, *Lipschitz spaces, smoothness of functions, and approximation theory*, Expositiones Mathematicae **3** (1983), 193–260.
6. A.P. Morse, *The behavior of a function on its critical set*, Ann. of Math. **40**, 62–70.
7. A. Norton, *A critical set with non-null image has large Hausdorff dimension*, Trans. Amer. Math. Soc. **296** (1986), 367–376.
8. ———, *Functions not constant on fractal quasi-arcs of critical points*, Proc. Amer. Math. Soc. **106** (1989), 397–406.
9. A. Norton and D. Sullivan, *Wandering domains and invariant conformal structures for mappings of the 2-torus*, preprint.
10. A. Sard, *The measure of the critical values of differentiable maps*, Bull. Amer. Math. Soc. **48** (1942), 883–890.
11. E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, 1970.
12. D. Varberg, *On differentiable transformations in R^n*, Amer. Math. Monthly **73** (1966), 111–114.
13. H. Whitney, *A function not constant on a connected set of critical points*, Duke Math. J. **1** (1935), 514–517.