# The Generalized Morse-Sard Theorem

Andrew McLennan*

**Abstract**

The Morse-Sard theorem gives conditions under which the set of critical values of a function between Euclidean spaces has Lebesgue measure zero. Over the years the result has been extended and strengthened in various ways. We present a result, along with a simple proof, that subsumes many of these generalizations. We also review methods of constructing examples showing that differentiability hypotheses cannot be weakened, and we construct a complete set of examples for our result.

**Keywords:** Morse-Sard theorem, Hausdorff measure, Hölder continuity, Whitney extension theorem.

## 1 Introduction

In its classic form the Morse-Sard theorem asserts that if $U \subset \mathbb{R}^m$ is open, $f\colon U \to \mathbb{R}^n$ is $C^{\max\{m-n+1,1\}}$, and $C_f = \{\, c \in U : \text{rank } Df(c) < n \,\}$ is the set of *critical points* of $f$ then the set $f(C_f)$ of *critical values* of $f$ has Lebesgue measure zero. This is one of the key results of $20^{\text{th}}$ century analysis. It is an essential foundation of the differential approach to the topology of manifolds. (This perspective is described beautifully by Milnor [33].) Among many other things, the Morse-Sard theorem has found many important applications in economic theory (e.g., [11, 26, 30]).

Over the years the result has been generalized in various directions. The purpose of this note is to present a result that incorporates most of these generalizations (except those related to Sobolev spaces) along with a proof that is (except for the proof of the final refinement) not more complex than other proofs in the literature. In the remainder of this section we state the result, give a brief overview of its historical development, and outline the remainder of the paper.

Let $E$ be $m$-dimensional Euclidean space, let $Y$ be a normed space, and fix an open $U \subset E$ and a function $f : U \to Y$. We endow the space of symmetric multilinear functions $\mu\colon E^i \to Y$ with the norm $|\mu| = \max_{|v_1|,\ldots,|v_i| \le 1} |\mu(v_1, \ldots v_i)|$. As always, if $k$ is a nonnegative integer, $f$ is $C^k$ if there are symmetric multilinear functions $D^i f(x)\colon E^i \to Y$ for $i =$

*School of Economics, The University of Queensland, Level 6 Colin Clark Building, St Lucia QLD 4072, Australia, a.mclennan@economics.uq.edu.au, September 21, 2020.

$0, \ldots, k$, which vary continuously with $x$, such that for each $x \in U$ and $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f(x+h) - \sum_{i=0}^{k} \tfrac{1}{i!} D^i f(x)(h, \ldots, h)| \le \varepsilon |h|^k$$

for all $h \in E$ such that $x + h \in U$ and $|h| < \delta$. Of course $D^0 f$ is simply $f$ itself, and we always write $Df$ in place of $D^1 f$. We always assume that $f$ is at least $C^1$. If $\varrho$ is a nonnegative integer, a point $c \in U$ is *$\varrho$-critical* for $f$ if $\operatorname{rank} Df(c) \le \varrho$; let $C_{f,\varrho}$ be the set of such points.

For $\alpha \in [0,1]$, $f$ is $C^{k,\alpha}$ ($C^{k,\alpha+}$) if it is $C^k$ and $D^k f$ satisfies the Hölder continuity condition that for each compact $K \subset U$ there is a constant $M_K > 0$ (a nondecreasing continuous $\epsilon_K \colon \mathbb{R} \to \mathbb{R}$ with $\epsilon_K(0) = 0$) such that

$$|D^k f(x) - D^k f(y)| \le M_K |x - y|^\alpha \quad (\le \epsilon_K(|x-y|)|x-y|^\alpha)$$

for all $x, y \in K$. Note that $C^{k,0+}$ is the same thing as $C^k$, and if $f$ is $C^{k,\alpha}$, then it is $C^{k,\alpha'+}$ for any $\alpha' < \alpha$. For a general $S \subset E$ a function $f \colon S \to Y$ is, by definition, $C^{k,\alpha}$ ($C^{k,\alpha+}$) if it has a $C^{k,\alpha}$ ($C^{k,\alpha+}$) extension $\tilde{f} \colon U \to Y$ to an open $U \subset E$ containing $S$.

For a subset $S \subset X$ of a metric space $(X, d)$ and positive numbers $\tau$ and $\beta$, let

$$\mathcal{H}^\tau_\beta(S) = \inf \sum_{i=1}^{\infty} (\operatorname{diam} T_i)^\tau$$

where $\operatorname{diam} T_i = \sup_{x,y \in T_i} d(x,y)$ and the infimum is over all countable collections $\{T_i\}$ of subsets of $X$ such that $\operatorname{diam} T_i < \beta$ for all $i$ and $S \subset \bigcup_i T_i$. Clearly $\mathcal{H}^\tau_\beta(S)$ is a nonincreasing function of $\beta$, and the $\tau$-dimensional *Hausdorff measure*[1] of $S$ is

$$\mathcal{H}^\tau(S) = \lim_{\beta \to 0} \mathcal{H}^\tau_\beta(S) \in [0, \infty].$$

Let $\mathcal{H}^0(S)$ be the cardinality of $S$ if $S$ is finite, and otherwise let $\mathcal{H}^0(S) = \infty$. If $\tau' > \tau > 0$, then $\mathcal{H}^{\tau'}_\beta(S) \le \beta^{\tau' - \tau} \mathcal{H}^\tau_\beta(S)$ for all $\beta$, so, for $\tau' > \tau \ge 0$, $\mathcal{H}^\tau(S) < \infty$ implies $\mathcal{H}^{\tau'}(S) = 0$ and $\mathcal{H}^{\tau'}(S) > 0$ implies $\mathcal{H}^\tau(S) = \infty$. The *Hausdorff dimension* of $S$ is the supremum of the set of $\tau$ such that $\mathcal{H}^\tau(S) > 0$. We say that $S$ is *$\tau$-null* if $\mathcal{H}^\tau(S) = 0$, *$\tau$-finite* if $\mathcal{H}^\tau(S) < \infty$, and *$\tau$-sigmafinite* if it is a countable union of $\tau$-finite sets. Note that if $Y$ is a second metric space, $f \colon X \to Y$ is Lipschitz, and $S$ is $\tau$-null (finite, sigmafinite) then $f(S)$ is $\tau$-null (finite, sigmafinite).

Throughout we work with a fixed integer $r$ such that $0 \le r < m$, a fixed integer $k \ge 1$, and a fixed number $\alpha \in [0,1]$. For $s \ge r$ let

$$d(s) = r + \frac{s - r}{k + \alpha}.$$

Our goal is to prove:

[1] Some authors define the $\tau$-dimensional Hausdorff measure of $S$ to be $\lim_{\beta \to 0} 2^{-\tau} \Gamma(\frac{1}{2})^\tau / \Gamma(\frac{\tau}{2} + 1) \mathcal{H}^\tau_\beta(S)$, which agrees with Lebesgue measure on $\mathbb{R}^\tau$ when $\tau$ is an integer, but for our purposes this is irrelevant.

**Theorem 1.** *If $f$ is $C^{k,\alpha}$ and $S \subset C_{f,r}$ is $s$-sigmafinite for some $s \geq r$, then $f(S)$ is $d(s)$-sigmafinite. If, in addition, either a) $S$ is $s$-null, b) $s > r$ and $f$ is $C^{k,\alpha+}$, or c) $s = m$, then $f(S)$ is $d(s)$-null.*

Assertions (a) and (c) were obtained by Bates and Moreira [4, 34]. Theorem 1 as a whole is a special case of the results of Ferone, Korobkov, and Roviello [19, 21], which contain the strongest and most general theorems of Morse–Sard type for smooth, Hölder, and Sobolev spaces. An innovation of Theorem 1 is the extension of some results (beyond those in [17]) to infinite dimensional range.

To see that Theorem 1 implies the classic form of the result suppose that $Y = \mathbb{R}^n$ and $f$ is $C^k$, where $k \geq \max\{m - n + 1, 1\}$. If $n - 1 < m$, then, since $f$ is $C^{k,0+}$ and $C_{f,n-1}$ is $m$-sigmafinite, either b) or c) of Theorem 1 implies that $f(C_{f,n-1})$ is $(n-1+(m-n+1)/k)$-null, so it is $n$-null. If $n > m$, then $C_{f,n-1} = U$ and a simple argument (along the lines of the proof of Proposition 1 below) implies that $f(U)$ is $m$-sigmafinite and consequently $n$-null. In either case it follows that $f(C_{f,n-1})$ has Lebesgue measure zero because $\mathcal{H}^n$ agrees (up to the normalizing constant) with Lebesgue outer measure on $\mathbb{R}^n$ ([15],[44]).

After preliminary results of Brown [8] and Marston Morse and Sard [37] the classic form of the result was proved in the case $n = 1$ by Anthony Morse [36], then in general by Sard [43], who also established that if $Y = \mathbb{R}^n$, $f$ is $C^k$, and $n \geq r + (m - r)/k$, then $f(C_{f,r})$ has Lebesgue measure zero. (Simple proofs of Theorem 1 for the case $Y = E$, $r = m - 1$, $k = 1$, $\alpha = 0$, and $s = m$ are given by Spivak [48, p. 72] and Constantin [9].) Sard [45] introduced the idea of studying the images of $s$-finite sets $S$ when $s < m$, and in [46] began the consideration of an infinite dimensional range. A jumble of partial results was unified and generalized by Federer [17, 3.4.3]. The present paper originated in an attempt to understand Federer's proof by simplifying and clarifying it, and Section 4 retains elements of his approach. At the same time as Federer's announcement [18] (without proof) of his result Dubovitskiĭ [14, Theorem 1] published a result that does not consider an infinite dimensional range, but is actually stronger than Federer's insofar as it considers images of $s$-finite sets when $s < \dim E$. (More precisely, it is Theorem 1 for $C^k$ functions and finite dimensional range.)

Hölder conditions on the top derivative were studied by Kaufman [27] (the case $r = 0$ and $s = m = 1$) Kučera [32] (the case $r = 0$ and $s = m$) and (independently) Norton [38]. The result above incorporates a suggested ([38, p. 369]) strengthening of Norton's main result. For $C^{k,\alpha}$ functions this strengthening was established by Bates and Moreira [4]. This strengthening for $C^{k,\alpha+}$ functions (that is, b)) is an original contribution of this paper.

Prior work related to c) assumes that $Y = \mathbb{R}^n$ is finite dimensional. Bates [2] established c) when $n = d(m)$, thereby (due to the case $r = n - 1$) weakening the differentiability requirement of the classic Morse-Sard theorem when $m > n$ from $C^{m-n+1}$ to $C^{m-n,1}$. Norton [40] weakened the requirement still further, to $f$ being $C^{m-n}$ and $D^{m-n}f$ being Zygmund, which is a weaker condition than Lipschitz. Moreira [34] established c) in general, except that he assumed a finite dimensional range, but we will see that the main ideas in his argument do not depend on this assumption.

Bates and Moreira introduce certain refinements in [3] and [4]. In [4] they assert that examples in [34] can easily be modified to show that their Theorem 1 is quite sharp for finite dimensional domains, and they, and also Kupka [31], give examples severely limiting possibilities for extension to infinite dimensional domains. In Section 7 we give a complete system of examples showing that the differentiability hypotheses of Theorem 1 are tight. So, in spite of a continuing evolution over several decades, there is perhaps some justification for describing Theorem 1 as "the" generalized Morse-Sard theorem, at least within some range.

Until recently most proofs of Morse-Sard type results used Morse decomposition, as we do here. This technique applies the implicit function theorem repeatedly in a double induction over dimension and degree of differentiability to produce a countable covering of the critical set by submanifolds of the domain, such that within each submanifold all critical points have the same lowest order of nonzero partial derivative and the same rank, and constitute a further submanifold. Insofar as this construction is cumbersome, it would be desirable to avoid it, but doing so requires methods that can analyze critical points of multiple types. Using metric entropy, Yomdin [52] developed quantitative measures of near criticality that provided an upper bound on the measure of the image of the near critical set, from which versions of the Morse-Sard theorem follow. Moreira and Ruas [35] use the curve selection theorem of semi-algebraic geometry to prove an inequality of Bochnak and Lojasiewicz, and in turn use this to prove that if $Y = \mathbb{R}$, $f$ is $C^k$, $x \in C_{f,0}$, and $\{x_i\}$ is a sequence in $C_{f,0} \setminus \{x\}$ converging to $x$, then $\lim |f(x_i) - f(x)|/|x_i - x|^k = 0$. Using the Vitali covering theorem, they then prove the Morse theorem (that is, the Morse-Sard theorem for one dimensional range) and then use Fubini's theorem to pass to the classical Morse-Sard theorem.

The first version of the Morse-Sard theorem for Sobolev spaces is due to de Pascale [10]. His argument uses a Morse decomposition, as do the proofs of Bojarski, Hajłasz, and Strzelecki [5] and Figalli [22]. Roughly between 2010 and 2012, Alberti and his coworkers [1], Bourgain, Korobkov, and Kristensen [6, 7, 28], and van der Putten [41, 42] independently developed results for Sobolev spaces and spaces of functions of bounded variation with proofs that do not use a Morse decomposition, but instead apply versions of the Luzin $N$ property. Among other things, some of these results allow functions that can fail to be $C^1$. This line of research continues to be quite active ([19, 20, 21, 24, 25, 29]) with [19] and [21] providing the most general results that contain others in this literature as special cases. Of these, [19] deserves special mention, as it extends Theorem 1 to Sobolev spaces.

In [12] Dubovitskiĭ proved the following result: if $n$ and $k$ are positive integers, $Y = \mathbb{R}^n$, $\nu = m - n - k + 1$, and $f$ is $C^k$, then $\mathcal{H}^\nu(C_{f,n-1} \cap f^{-1}(y)) = 0$ for almost all $y \in \mathbb{R}^n$. Dubovitsiĭ's article has never been translated, and went largely unnoticed for many years, but this approach has strongly influenced recent research ([6, 7, 19, 20, 24, 25, 28]). In particular, [19] gives the following version of it for the classical case, thereby refining Theorem 1: Suppose that $Y = \mathbb{R}^n$, $f$ is $C^{k,\alpha}$, $s \in [r, m]$, $q \in [r, d(s)]$, $\mu = (d(s) - q)(k + \alpha)$, and $S \subset C_{f,r}$. If $S$ is $s$-null, or if $S$ is $s$-sigmafinite and either $s = m$ or $s > r$ and $f$ is $C^{k,\alpha+}$, then $\mathcal{H}^\mu(S \cap f^{-1}(y)) = 0$ for $\mathcal{H}^q$-almost all $y \in \mathbb{R}^n$.

Finally we should mention that Smale [47] has given a version of the Morse-Sard theorem for functions whose domain and range are both infinite dimensional.

Each individual element of Theorem 1 is present in prior literature, but no single predecessor incorporates all of them simultaneously, and Theorem 1 has an original finding. But certainly this paper's more striking contribution is the proof, which is simple and unified. Section 2 presents a change of coordinates that simplifies many aspects of the subsequent argument. After that the proof has four main elements. Propositions 1 and 2 are measurements that will be used to establish all of the assertions aside from c). The central computation of the proof of Proposition 2 was adapted from the proof of [14, Lemma 8]; similar calculations have appeared in other sources. Section 4 presents what is, in effect, a version of Taylor's theorem for the objects coming from the Morse decomposition of the critical set; our argument follows [17, 3.4.2]. The third main element is the Morse decomposition itself (Proposition 4) which we prove in the usual way. The fourth main element is the proof of c) (Proposition 6). This is the most complicated part of the paper. It uses ideas from Moreira's [34] argument (see also [3]) but transports them to a more elementary framework similar to [2], which simplifies various aspects.

Section 7 briefly describes methods for constructing examples, used to show that various hypotheses cannot be weakened, that were developed by Whitney [51], Dubovitskiĭ [12, 13, 14], and Federer [17, 3.4.4]. The last method is used to construct a complete suite of examples showing that the differentiability hypotheses of Theorem 1 cannot be weakened.

Appendix A presents a version of the Whitney extension theorem [50] for $C^{k,\alpha}$ and $C^{k,\alpha+}$ functions. This result was established by Norton [39], and was discussed earlier by Stein [49]. Appendix B is a slight expansion of [38, Appendix] that presents versions of basic calculus results for $C^{k,\alpha}$ and $C^{k,\alpha+}$ functions.

# 2 Straightening

The discussion in this section considers only $C^{k,\alpha}$ objects, but it is completely correct if we simply substitute $C^{k,\alpha+}$ everywhere, so it should be understood as pertaining equally to both differentiability classes.

We begin by showing that a simplifying change of coordinates is possible. Throughout the remainder we work with a fixed $(m-r)$-dimensional linear subspace $Z \subset E$.

**Lemma 1.** *If $f$ is $C^{k,\alpha}$, $x_0 \in U$, and $\ker Df(x_0) = Z$, then there is an open neighborhood $V \subset U$ of $x_0$, an open $\tilde{V} \subset E$, and a $C^{k,\alpha}$ diffeomorphism $g\colon V \to \tilde{V}$ such that $\tilde{f} = f \circ g^{-1}$ is $C^{k,\alpha}$ and for each $\tilde{x} \in \tilde{V}$, $\operatorname{rank} D\tilde{f}(\tilde{x}) \le r$ if and only if $\ker D\tilde{f}(\tilde{x}) = Z$.*

*Proof.* Let $W$ be a linear subspace of $E$ such that $W \cap Z = \{0\}$ and $W + Z = E$, and let $p_W\colon w + z \mapsto w$ and $p_Z\colon w + z \mapsto z$ be the linear projections. The Hahn-Banach theorem gives a continuous linear $L\colon Y \to W$ that extends $(Df(x_0)|_W)^{-1}$. Let $g\colon U \to E$ be the function $g(x) = L(f(x)) + p_Z(x)$. Since $Dg(x_0)|_W$ and $Dg(x_0)|_Z$ are the respective identities, the inverse function theorem (Lemma B3) gives a neighborhood $V \subset U$ of $x_0$

such that $g|_V$ is invertible with $C^{k,\alpha}$ inverse. Let $\tilde{V} = g(V)$ and $\tilde{f} = f \circ (g|_V)^{-1}$. (This is $C^{k,\alpha}$ by Lemma B2.)

Fix $x \in V$, and let $\tilde{x} = g(x)$, noting that $\tilde{f}(\tilde{x}) = f(x)$. We have

$$p_W(\tilde{x}) = p_W(L(f(x)) + p_Z(x)) = p_W(L(f(x))) = L(f(x)) = L(\tilde{f}(\tilde{x})),$$

so $L \circ \tilde{f} = p_W|_{\tilde{V}}$. Therefore $L \circ D\tilde{f}(\tilde{x}) = D(L \circ \tilde{f})(\tilde{x}) = p_W$, and consequently $\operatorname{rank} D\tilde{f}(\tilde{x}) \geq \operatorname{rank} L \circ D\tilde{f}(\tilde{x}) = r$. In addition $D\tilde{f}(\tilde{x})W \cap \ker L = \{0\}$ and $D\tilde{f}(\tilde{x})Z \subset \ker L$. If $\operatorname{rank} D\tilde{f}(\tilde{x}) = r$, then $D\tilde{f}(\tilde{x})E = D\tilde{f}(\tilde{x})W$, so $D\tilde{f}(\tilde{x})Z \subset D\tilde{f}(\tilde{x})W \cap \ker L = \{0\}$. $\square$

Suppose $X$ is a metric space, $S \subset X$, and each $x \in S$ has a neighborhood $U$ such that $S \cap U$ is $\tau$-null. If $X$ is separable with countable dense subset $D$, then each $x \in S$ is contained in a $\tau$-null ball of rational radius centered at a point in $D$. Since the set of such balls is countable, $S$ is $\tau$-null. The same line of reasoning applies to sets that are $\tau$-sigmafinite, so to prove Theorem 1 it suffices to show that each $x_0 \in S$ has an open neighborhood $V$ and a compact neighborhood $K \subset V$ such that $f(S \cap K)$ is $d(s)$-null or $d(s)$-sigmafinite, as required by the relevant assertion. Suppose $f$, $V$, $g$, $\tilde{V}$, and $\tilde{f}$ are as in Lemma 1, $K \subset V$ is a compact neighborhood of $x_0$, $S \subset V$, $\tilde{K} = g(K)$, and $\tilde{S} = g(S)$. Since $g|_K$ and its inverse are $C^1$, they are locally Lipschitz, hence Lipschitz[2], so $S \cap K$ is $s$-null (sigmafinite) if and only if $\tilde{S} \cap \tilde{K}$ is $s$-null (sigmafinite). Since $d(s)$ is an increasing function of $r$ it suffices to prove Theorem 1 with the condition $S \subset C_{f,r}$ replaced by $S \subset C_{f,r} \setminus C_{f,r-1}$. Since $\tilde{S} \subset \{\, \tilde{c} \in \tilde{V} : \ker D\tilde{f}(\tilde{c}) = Z \,\}$, the upshot of this discussion is that it suffices to prove Theorem 1 with the hypothesis $S \subset C_{f,r}$ replaced by $S \subset \{\, c \in U : \ker Df(c) = Z \,\}$.

For us a $C^{k,\alpha}$ *manifold* (with one coordinate chart) is a set $N \subset E$ such that for some integer $\mu_N$ there is a $\mu_N$-dimensional linear subspace $W_N \subset E$, an open $U_N \subset W$, and a $C^{k,\alpha}$ diffeomorphism $\psi\colon U \to N$. In this circumstance the *tangent space* $T_xN$ of $N$ at $x \in N$ is the image of $D\psi_N(\psi_N^{-1}(x))$. The following definition provides useful additional objects and restrictions in relation to $Z$. A tuple $(N, \mu_N, W_N, U_N, \psi_N, Z_N, A_N, \rho_N)$ is a *full featured $C^{k,\alpha}$ manifold* if:

(a) $W_N$ is a $\mu_N$-dimensional linear subspace of $E$, $U_N \subset W_N$ is open, and $\psi_N\colon U_N \to N \subset E$ is a $C^{k,\alpha}$ diffeomorphism.

(b) $T_xN + Z = E$ for each $x \in N$. (This implies that $\mu_N \geq r$.)

(c) $Z_N$ is a $(\mu_N - r)$-dimensional linear subspace of $W_N$ such that:

  (i) For all $\xi \in U_N$ and $\nu \in W_N$, $\nu \in Z_N$ if and only if $D\psi_N(\xi)\nu \in Z$.

  (ii) For all $\xi, \xi' \in U_N$, $\xi - \xi' \in Z_N$ if and only if $\psi_N(\xi) - \psi_N(\xi') \in Z$.

(d) $A_N \subset E$ is an open neighborhood of $N$, and $\rho_N\colon A_N \to N$ is a $C^{k,\alpha}$ retraction such that $y - \rho_N(y) \in Z$ for all $y \in A_N$.

[2]Suppose $X$ and $Y$ are metric spaces, $X$ is compact, and $f\colon X \to Y$ is locally Lipschitz. If $d(f(x_n), f(x'_n))/d(x_n, x'_n) \to \infty$, then $d(x_n, x'_n) \to 0$ because $\operatorname{diam} f(X) < \infty$, so there is a common limit point $x$, and $f$ is Lipschitz in some neighborhood of $x$, contrary to assumption. Thus $f$ is Lipschitz.

When speaking of such an object, $N$ may refer either to the entire tuple or to the manifold itself, with the correct interpretation to be inferred from context.

# 3 Measurements

Fix a full featured $C^1$ manifold $N$. This section gives two results that measure $f(S)$ for $S \subset N$ under certain circumstances. The first is simple, and illustrates the methods that will be applied in the more substantial results that follow. The second may be regarded as the driving force of Theorem 1.

**Proposition 1.** *If $\mu_N = r$ and $\tau \geq r$, then $f(N)$ is $\tau$-sigmafinite, and if $\tau > r$, then $f(N)$ is $\tau$-null.*

*Proof.* If $r = 0$, then $M_N = \{0\}$, so $N$ and $f(N)$ are singletons. Otherwise $U_N$ can be covered by a countable collection of cubes, and it suffices to show that $f(\psi_N(C))$ is $\tau$-finite, and $\tau$-null if $\tau > r$, when $C \subset U_N$ is a compact cube, say of sidelength $\ell$. Since $C$ is compact and $f \circ \psi_N$ is $C^1$, $f \circ \psi_N|_C$ is Lipschitz; let $L$ be a Lipschitz constant. For a positive integer $P$, $C$ can be covered by $P^r$ subcubes $C_i$ of sidelength $\ell/P$. We have $\operatorname{diam} f(\psi_N(C_i)) \leq L \cdot \operatorname{diam} C_i = L\ell\sqrt{r}/P$. Of course $L\ell\sqrt{r}/P \to 0$ as $P \to \infty$, and

$$\mathcal{H}^\tau_{L\ell\sqrt{r}/P}(f(\psi_N(C))) \leq P^r(L\ell\sqrt{r}/P)^\tau = (L\ell\sqrt{r})^\tau P^{r-\tau}. \qquad \square$$

We now introduce a function from [17, 3.4.2]. If $K \subset N$ is compact, $0 \leq \kappa < \infty$, and $\delta > 0$, let

$$\eta(N, K, f, \kappa, \delta) = \sup_{c \in K,\, x \in N,\, 0 \neq x - c \in Z,\, |x-c| \leq \delta} \frac{|f(x) - f(c)|}{|x - c|^\kappa}.$$

For this definition, if there are no $c$ and $x$ with the required properties we convene that the defined number is zero. Since $\eta(N, K, f, \kappa, \delta)$ is a nondecreasing functions of $\delta$, its limit as $\delta \to 0$ is well defined. We denote this limit by $\eta(N, K, f, \kappa)$.

The next result will be applied with $\kappa = k + \alpha$ and $\omega = \frac{s-r}{k+\alpha}$. The proof below is derived from the proof of [14, Lemma 8].

**Proposition 2.** *Suppose that $\kappa \geq 1$ and $\omega \geq 0$ are real numbers, $r + \omega > 0$, $\mu_N > r$, and $K \subset N$ is compact. Let $T = \eta(N, K, f, \kappa)$. If $T < \infty$ and $S \subset K$ is $(r + \omega\kappa)$-finite, then $f(S)$ is $(r + \omega)$-finite. If, in addition, either $S$ is $(r + \omega\kappa)$-null or $\omega > 0$ and $T = 0$, then $f(S)$ is $(r + \omega)$-null.*

*Proof.* Let $f' = f \circ \psi_N$, $K' = \psi_N^{-1}(K)$, and $S' = \psi_N^{-1}(S)$. The restriction of $\psi_N$ to a compact neighborhood of $K'$ is Lipschitz, because $\psi_N$ is $C^1$, so $\eta(U_N, K', f', \kappa) < \infty$. Thus it suffices to prove the assertion with $N$, $K$, $f$, and $S$ replaced by $U_N$, $K'$, $f'$, and $S'$, so we may assume that $N$ is an open subset of $W_N$, in which case (c) reduces to $Z_N = W_N \cap Z$.

Let $\tilde{K} \subset U_N$ be a compact neighborhood of $K$. Since $f$ is $C^1$, $f|_{\tilde{K}}$ is Lipschitz; let $L$ be a Lipschitz constant. Fix $\varepsilon > 0$ such that $\varepsilon\sqrt{\mu_N - r} < L\sqrt{r}$. Fix $\delta > 0$ small enough that

$\tilde{K}$ contains the ball of radius $\delta$ around $K$ and $\eta(N, K, f, \kappa, \delta) < T + \varepsilon$. Fix $\theta > 0$ small enough that $\mathcal{H}_\theta^{r+\omega\kappa}(S) < \mathcal{H}^{r+\omega\kappa}(S) + \varepsilon$, $\sqrt{\mu_N}\theta < \delta$, and $\varepsilon(2\sqrt{\mu_N - r})^\kappa \theta^{\kappa-1} < 2L\sqrt{r}$.

Let $\{S_i\}$ be a countable cover of $S$ by sets $S_i$ of diameter $d_i < \theta$ such that $\sum_i d_i^{r+\omega\kappa} < \mathcal{H}^{r+\omega\kappa}(S) + \varepsilon$. For each $i$ let $C_i = A_i \times B_i$ be a cube of sidelength $2d_i$ centered at a point of $S_i$, where $B_i$ is a cube in $Z_N$ and $A_i$ is a cube in its orthogonal complement. Of course $S_i \subset C_i$, and $C_i \subset \tilde{K}$ because $\sqrt{\mu_N} d_i < \delta$. Let $D_i^A = 2L\sqrt{r} d_i$ and $D_i^B = 2\sqrt{\mu_N - r} d_i$. Let $G = 2^{r+\kappa\omega} r^{r/2} (\mu_N - r)^{\kappa\omega/2} L^r$, so that

$$\sum_i (D_i^A)^r (D_i^B)^{\kappa\omega} = G \sum_i d_i^{r+\kappa\omega} < G(\mathcal{H}^{r+\kappa\omega}(S) + \varepsilon).$$

Since

$$D_i^A / \varepsilon (D_i^B)^\kappa \geq 2L\sqrt{r} / \varepsilon(2\sqrt{\mu_N - r})^\kappa \theta^{\kappa-1} > 1$$

there is an integer $P_i$ such that

$$\frac{D_i^A}{(T+\varepsilon)(D_i^B)^\kappa} \leq P_i < \frac{2D_i^A}{(T+\varepsilon)(D_i^B)^\kappa}.$$

We cover $A_i$ with $P_i^r$ subcubes $A_{ij}$ of sidelength $2d_i/P_i$. We claim that

$$\text{diam } f((A_{ij} \times B_i) \cap S) \leq 2(T+\varepsilon)(D_i^B)^\kappa.$$

Suppose that $c, c' \in (A_{ij} \times B_i) \cap S$ where $c = (a, b)$ and $c' = (a', b')$, and let $x = (a, b')$. Then $|f(c') - f(x)| \leq L|a' - a| \leq D_i^A / P_i$. We have $|x - c| = |b' - b| \leq D_i^B$, and $b' - b \in Z$ and $|b - b'| \leq 2\sqrt{\mu_N - r} d_i < \delta$, so

$$|f(x) - f(c)| \leq (T+\varepsilon)|x - c|^\kappa \leq (T+\varepsilon)(D_i^B)^\kappa,$$

and thus, as claimed,

$$|f(c') - f(c)| \leq D_i^A / P_i + (T+\varepsilon)(D_i^B)^\kappa \leq 2(T+\varepsilon)(D_i^B)^\kappa.$$

Let $\beta = 2(T+\varepsilon)(2\sqrt{\mu_N - r}\theta)^\kappa$, so that diam $f((A_{ij} \times B_i) \cap S) < \beta$ for all $i$ and $j$. Since $P_i < 2D_i^A/(T+\varepsilon)(D_i^B)^\kappa$ we now have

$$\begin{aligned}\mathcal{H}_\beta^{r+\omega}(f(S)) &\leq \sum_{i,j} (\text{diam } f((A_{ij} \times B_i) \cap S))^{r+\omega} \leq \sum_i \left( \frac{2D_i^A}{(T+\varepsilon)(D_i^B)^\kappa} \right)^r (2(T+\varepsilon)(D_i^B)^\kappa)^{r+\omega} \\ &= 2^{2r+\omega}(T+\varepsilon)^\omega \sum_i (D_i^A)^r (D_i^B)^{\kappa\omega} < 2^{2r+\omega} G (T+\varepsilon)^\omega (\mathcal{H}^{r+\kappa\omega}(S) + \varepsilon).\end{aligned}$$

Since $\beta$ and $\varepsilon$ may be arbitrarily small, the desired assertions follow from this. □

# 4 Orders of Magnitude

The remainder of the proof of Theorem 1 is a matter of finding a countable cover of $C_{f,r}$ by compact sets to which Propositions 1 and 2 can be applied. The hypothesis on $\eta(N, K, f, \kappa)$

resembles the conclusion of Taylor's theorem, and this section's argument obtains this via an induction in which the desired conclusion at one degree of differentiability is integrated, using Lemmas 2 and 3 below, to obtain the desired conclusion at the next degree. This generalizes (from $C^k$ to $C^{k,\alpha}$ and $C^{k,\alpha+}$) the analysis in [17, 3.4.2].

We introduce a second function from [17, 3.4.2]. If $N$ is a full featured $C^1$ manifold, $K \subset N$ is compact, $0 \le \kappa < \infty$, and $\delta > 0$, let

$$\zeta(N, K, f, \kappa, \delta) = \sup_{\substack{c \in K,\, x \in N,\, 0 \ne x - c \in Z,\, |x-c| \le \delta, \\ z \in T_x N \cap Z,\, |z| \le 1}} \frac{|Df(x)z|}{|x-c|^\kappa}.$$

If there are no $c$, $x$, and $z$ with the required properties we convene that the defined number is zero. As before, $\zeta(N, K, f, \kappa, \delta)$ is a nondecreasing functions of $\delta$, and we denote its limit as $\delta \to 0$ by $\zeta(N, K, f, \kappa)$.

**Lemma 2.** *If $N$ is a full featured $C^1$ manifold, $K \subset N$ is compact, $\kappa \ge 1$, and $\zeta(N, K, f, \kappa - 1) < \infty$ $(= 0)$, then $\eta(N, K, f, \kappa) < \infty$ $(= 0)$.*

*Proof.* Let $\varepsilon > 0$ be small enough that $\tilde{K}_\varepsilon = \{ x \in E : \mathrm{dist}(x, K) \le \varepsilon \}$ is contained in $A_N$. Since $\rho_N$ is $C^1$ and $\tilde{K}_\varepsilon$ is compact, $\rho_N|_{\tilde{K}_\varepsilon}$ is Lipschitz; let $L$ be a Lipschitz constant. We will show that $\eta(N, K, f, \kappa) \le L^\kappa \zeta(N, K, f, \kappa - 1)$.

Suppose that $0 < \delta < \varepsilon$, $c \in K$, and $x \in N$, with $0 \ne x - c \in Z$ and $|x - c| \le \delta / L$. (If there are no such $c$ and $x$, then the claim is an automatic consequence of our conventions.) Let $\pi \colon [0,1] \to N$ be the $C^1$ path $\pi(t) = \rho_N(c + t(x - c))$. For each $t$ we have

$$\pi(t) - c = \big(\rho_N(c + t(x-c)) - c - t(x-c)\big) + t(x-c) \in Z,$$

so $\pi'(t) \in T_{\pi(t)} N \cap Z$. Evidently $|D\rho_N(x)y| \le L|y|$ for all $x \in \tilde{K}_\varepsilon$ and $y \in E$, so $|\pi'(t)| \le L|x - c|$ and $|\pi(t) - c| \le L|x - c| \le \delta$. Therefore

$$\big|Df(\pi(t))\pi'(t)/|\pi'(t)|\big| \le \zeta(N, K, f, \kappa - 1, \delta)|\pi(t) - c|^{\kappa-1} \le L^{\kappa-1}\zeta(N, K, f, \kappa - 1, \delta)|x - c|^{\kappa-1}.$$

We conclude that $\eta(N, K, f, \kappa, \delta/L) \le L^\kappa \zeta(N, K, f, \kappa - 1, \delta)$ because

$$\begin{aligned} |f(x) - f(c)| &\le \int_0^1 |(f \circ \pi)'(t)|\, dt \le \int_0^1 \big|Df(\pi(t))\pi'(t)/|\pi'(t)|\big| \cdot |\pi'(t)|\, dt \\ &\le L^\kappa \zeta(N, K, f, \kappa - 1, \delta)|x - c|^\kappa. \end{aligned}$$ □

While one would typically prove Taylor's theorem in a pleasantly structured setting, our lack of prior knowledge concerning the structure of $C_{f,r}$ requires us to work in a rather cumbersome structure. A $(k, \alpha)$-*flattening configuration* ($(k, \alpha+)$-*flattening configuration*) is a $k$-tuple $((N_1, K_1), \ldots, (N_k, K_k))$ such that:

(a) Each $N_\ell$ is a full featured $C^{\ell,\alpha}$ ($C^{\ell,\alpha+}$) manifold, and $N_1 \subset \cdots \subset N_k$.

(b) Each $K_\ell$ is a compact subset of $N_\ell$, and $K_1 \subset \cdots \subset K_k$.

(c) For each $\ell = 2, \dots, k$, if $\phi \colon N_\ell \to Y$ is a $C^{\ell,\alpha}$ ($C^{\ell,\alpha+}$) function that vanishes on $K_\ell$, then $D\phi(x)z = 0$ for all $x \in K_{\ell-1}$ and $z \in T_x N_\ell \cap Z$.

Fix such a configuration $((N_1, K_1), \dots, (N_k, K_k))$. For each $z \in Z$ let $\varphi_z \colon N_k \to Y$ be the function $\varphi_z(x) = Df(x)(D\rho_{N_k}(x)z) = D(f \circ \rho_{N_k})(x)z$.

**Lemma 3.** *Suppose that $Df(c)z = 0$ for all $c \in K_k$ and $z \in T_c N_k \cap Z$. If $\eta(N_1, K_1, \varphi_z, k - 1 + \alpha) < \infty$ $(= 0)$ for all $z \in Z$, then $\zeta(N_1, K_1, f, k - 1 + \alpha) < \infty$ $(= 0)$.*

*Proof.* If $c \in K_k$ and $z \in Z$, then $D\rho_N(c)z \in Z$ (because $\mathrm{Id}_{A_N} - \rho_N$ maps into $Z$) and of course $D\rho_N(c)z \in T_c N_k$, so $\varphi_z(c) = 0$. Let $e_1, \dots, e_{m-r}$ be an orthonormal basis of $Z$. If $c \in K_1$, $x \in N_1$, $x \neq c$, $x - c \in Z$, $z \in T_x N_1 \cap Z$, and $|z| \leq 1$, then $D\rho_{N_k}(x)z = z$ because $T_x N_1 \subset T_x N_k$ and $\rho_{N_k}|_{N_k}$ is the identity, so

$$|Df(x)z| = |Df(x)(D\rho_{N_k}(x)z)| = |\sum_i \langle z, e_i \rangle \varphi_{e_i}(x)| \leq |z| \sum_i |\varphi_{e_i}(x) - \varphi_{e_i}(c)|.$$

Dividing this by $|x - c|^{k-1+\alpha}$ reveals that

$$\zeta(N_1, K_1, f, k - 1 + \alpha) \leq \sum_i \eta(N_1, K_1, \varphi_{e_i}, k - 1 + \alpha). \qquad \square$$

**Proposition 3.** *If $f|_{N_k}$ is $C^{k,\alpha}$ ($C^{k,\alpha+}$) with $Df(c)z = 0$ for all $c \in K_k$ and $z \in T_c N_k \cap Z$, then $\eta(N_1, K_1, f, k + \alpha) < \infty$ $(= 0)$.*

*Proof.* In view of Lemma 2 it suffices to prove that $\zeta(N_1, K_1, f, k - 1 + \alpha) < \infty$ $(= 0)$, which we do by induction on $k$. If $k = 1$ and $K_1' \subset N_1$ is a compact neighborhood of $K_1$, then, since $D(f \circ \rho_{N_1})$ is $C^{0,\alpha}$ ($C^{0,\alpha+}$), there is an $M_{K'} > 0$ (a nondecreasing continuous $\epsilon_{K'} : \mathbb{R} \to \mathbb{R}$ with $\epsilon_{K'}(0) = 0$) such that

$$|D(f \circ \rho_{N_1})(x) - D(f \circ \rho_{N_1})(c)| \leq M_{K'} |x - c|^\alpha \ (\leq \epsilon_{K'}(|x - c|) |x - c|^\alpha)$$

for all $c \in K_1$ and $x \in K_1'$. If $c \in K_1$, $x \in K_1'$, $x - c \in Z$, and $z \in T_x N_1 \cap Z$ with $|z| \leq 1$, then, as above, $D\rho_{N_1}(x)z = z$ and $D\rho_{N_1}(c)z \in T_c N_1 \cap Z$, so

$$|Df(x)z| = |Df(x)(D\rho_{N_1}(x)z) - Df(c)(D\rho_{N_1}(c)z)| \leq |D(f \circ \rho_{N_1})(x) - D(f \circ \rho_{N_1})(c)|.$$

Suppose that $k \geq 2$ and we have proved the desired inequality with $k$ replaced by $k - 1$. As a composition of $C^{k-1,\alpha}$ ($C^{k-1,\alpha+}$) functions, $\varphi_z$ is $C^{k-1,\alpha}$ ($C^{k-1,\alpha+}$). Since $\varphi_z(c) = 0$ for all $c \in K_k$, $D\varphi_z(c)z = 0$ for all $c \in K_{k-1}$ and $z \in T_c N_k \cap Z$. The induction hypothesis (the analysis above is valid with $\varphi_z$ in place of $f$) gives $\zeta(N_1, K_1, \varphi_z, k - 2 + \alpha) < \infty$ $(= 0)$. Now Lemma 2 yields $\eta(N_1, K_1, \varphi_z, k - 1 + \alpha) < \infty$ $(= 0)$, after which Lemma 3 gives the desired result. $\square$

# 5 Morse Decomposition

The discussion in this section, up to Proposition 5, is completely correct with $C^{k,\alpha}$ replaced throughout by $C^{k,\alpha+}$, and should be understood as pertaining to both differentiability classes. Let $M$ be a full featured $C^{k,\alpha}$ manifold, and let $C \subset M$ be closed. A $C^{k,\alpha}$ *Morse decomposition* of $C$ is a countable collection $\{((N_1^j, K_1^j), \dots, (N_k^j, K_k^j))\}$ of $(k,\alpha)$-flattening configurations with $N_k^j \subset M$ for all $j$ and $\bigcup_j K_1^j = C$. Let $Q$ be the set of $x \in C$ such that $D\phi(x)z = 0$ for all $z \in T_xM \cap Z$ whenever $\phi\colon M \to Y$ is a $C^{k,\alpha}$ function that vanishes on $C$. If $x \in C \setminus Q$, $\phi\colon M \to Y$ is a $C^{k,\alpha}$ function that vanishes on $C$, $z \in T_xM \cap Z$, and $D\phi(x)z \neq 0$, then a neighborhood of $x$ in $C$ can be "peeled off" into a $(\mu_M - 1)$-dimensional submanifold of $M$ be applying the implicit function theorem, and this will allow a Morse decomposition to be constructed by a double induction on $\mu_M$ and $k$.

First we need to show that we can require the result of the application of the implicit function theorem to be full featured. One of these features can be fabricated from scratch.

**Lemma 4.** *If $N$ is a $C^{k,\alpha}$ manifold, $x_0 \in N$, and $T_{x_0}N + Z = E$, then there is an open neighborhood $N' \subset N$ of $x_0$ for which there is an open neighborhood $A \subset E$ of $N'$ and a $C^{k,\alpha}$ retraction $\rho\colon A \to N'$ such that $y - \rho(y) \in Z$ for all $y \in A$.*

*Proof.* Let $Z'$ be an $(m - \mu_N)$-dimensional linear subspace of $Z$ such that $T_{x_0}N + Z' = E$. The function $(x', z') \mapsto x' + z'$ from $N \times Z'$ to $E$ is $C^{k,\alpha}$, and its derivative at $(x_0, 0)$ is nonsingular, so Lemma B3 gives a neighborhood $B$ of $(x_0, 0)$ and a $C^{k,\alpha}$ diffeomorphism between $B$ and a neighborhood $A \subset E$ of $x_0$. Let the inverse of this diffeomorphism be $(\rho, \sigma)\colon A \to N \times Z'$, and let $N' = \{\, x \in N : (x, 0) \in B \,\}$. We now replace $A$ with $\rho^{-1}(N')$ and $\rho$ with its restriction to this set. □

**Lemma 5.** *Let $N$ be a full featured $C^{k,\alpha}$ manifold. If $\phi\colon N \to Y$ is $C^{k,\alpha}$, $x_0 \in \phi^{-1}(0)$, $z \in T_{x_0}N \cap Z$, and $D\phi(x_0)z \neq 0$, then there is a full featured $(\mu - 1)$-dimensional $C^{k,\alpha}$ manifold $N'$ that contains a neighborhood of $x_0$ in $\phi^{-1}(0)$.*

*Proof.* The Hahn-Banach theorem gives a continuous linear $L\colon Y \to \mathbb{R}$ with $L(D\phi(x_0)z) \neq 0$. Let $\phi' = L \circ \phi$. Since $D\phi'(x_0)z = (L \circ D\phi(x_0))z \neq 0$ and $\phi^{-1}(0) \subset {\phi'}^{-1}(0)$, it suffices to prove the claim with $\phi$ replaced by $\phi'$, so we may assume that $Y = \mathbb{R}$. Let $\tilde{\phi} = \phi \circ \psi_N\colon U_N \to \mathbb{R}$.

Let $\xi_0 = \psi_N^{-1}(x_0)$ and $\zeta = D(\psi_N^{-1})(x_0)z$. Let $\mu_{N'} = \mu_N - 1$, and let $W_{N'}$ be a $\mu_{N'}$-dimensional linear subspace of $W_N$ that does not contain $\zeta$. Since $D\tilde{\phi}(x_0)\zeta = D\phi(x_0)z \neq 0$, the $C^{k,\alpha}$ implicit function theorem (Lemma B4) gives an open $U_{N'} \subset W_{N'}$ and a $C^{k,\alpha}$ function $g\colon U_{N'} \to \mathbb{R}$ such that $\{\, \xi + g(\xi)\zeta : \xi \in U_{N'} \,\}$ is a neighborhood of $\xi_0$ in $\tilde{\phi}^{-1}(0)$. Let $\psi_{N'}\colon U_{N'} \to N$ be the function

$$\psi_{N'}(\xi) = \psi_N(\xi + g(\xi)\zeta).$$

Let $N' = \psi_{N'}(U_{N'})$; of course $N'$ is a neighborhood of $x_0$ in $\phi^{-1}(0)$.

By construction $T_{x_0}N'$ and $z$ span $T_{x_0}N$, and $T_{x_0}N$ and $Z$ span $E$, so $T_{x_0}N'$ and $Z$ span $E$, and if $U_{N'}$ is a small enough neighborhood of $\xi_0$, then by continuity $T_{x'}N'$ and $Z$ span

$E$ for all $x' \in N'$. Thus $N'$, $W_{N'}$, $U_{N'}$, and $\psi_{N'}$ satisfy (a) and (b) of the definition of a full featured manifold.

Let $Z_{N'} = W_{N'} \cap Z_N$. By (i) of (c), $\zeta \in Z_N$, so $Z_{N'}$ has codimension 1 in $Z_N$ and is thus $(\mu_{N'} - r)$-dimensional. Consider $\xi \in U_{N'}$ and $\nu \in W_{N'}$. Since $N$ satisfies (i), $D\psi_N(\xi+g(\xi)\zeta)\nu \in Z$ if and only if $\nu \in Z_N$, and $D\psi_{N'}(\xi)\nu = D\psi_N(\xi+g(\xi)\zeta)\nu+(Dg(\xi)\nu)z$, so $D\psi_{N'}(\xi)\nu \in Z$ if and only if $\nu \in Z_{N'}$. Consider $\xi, \xi' \in U_{N'}$. Since $N$ satisfies (ii), $\psi_{N'}(\xi) - \psi_{N'}(\xi') \in Z$ if and only if $(\xi + g(\xi)\zeta) - (\xi' + g(\xi')\zeta) \in Z_N$, and this is the case if and only if $\xi - \xi' \in Z_{N'}$. Thus (c) holds.

Now Lemma 4 allows us to replace $N'$ (with related adjustments to $U_{N'}$ and $\psi_{N'}$) with an open neighborhood of $x_0$ for which there is an open neighborhood $A_{N'} \subset E$ of $N'$ and a $C^{k,\alpha}$ retraction $\rho_{N'} \colon A_{N'} \to N'$ such that $y - \rho(y) \in Z$ for all $y \in A_{N'}$. □

**Proposition 4.** *There is $C^{k,\alpha}$ Morse decomposition of $C$.*

*Proof.* A collection of Morse decompositions of the elements of a countable cover of $C$ by compact sets can be united to give a Morse decomposition of $C$, so we may assume that $C$ itself is compact. Note that (c) of the definition of a $(k, \alpha)$-flattening configuation holds vacuously when $k = 1$, and it also holds when $N_\ell$ is $r$-dimensional because $T_x N_\ell \cap Z = \{0\}$ for all $x \in N_\ell$. When $k = 1$, $\{((M, C))\}$ is a Morse decomposition of $C$. If $\mu_M = r$, then $Z_M = \{0\}$, so a suitable collection is given by $\{((M, K_1), \ldots, (M, K_k))\}$ where $K_1 = \cdots = K_k = C$. By induction it suffices to show that the claim holds for given $\mu_M > r$ and $k > 1$ if it has already been established with $(\mu_M, k)$ replaced by either $(\mu_M - 1, k)$ or $(\mu_M, k - 1)$.

Let $Q$ be as above. If $x \in C \setminus Q$, $\phi \colon M \to Y$ is a $C^{k,\alpha}$ function that vanishes on $C$, $z \in T_x M \cap Z$, and $D\phi(x)z \neq 0$, then Lemma 5 gives a full featured $(\mu_M - 1)$-dimensional $C^{k,\alpha}$ manifold containing a neighborhood of $x$ in $\phi^{-1}(0) \supset C$. Since $C \setminus Q$ is separable, there is a countable cover $\{N^i\}$ of $C \setminus Q$ by such manifolds. By induction on $\mu_M$, for each $i$ there a countable collection $\{((N^{i\ell}_1, K^{i\ell}_1), \ldots, (N^{i\ell}_k, K^{i\ell}_k))\}$ of $(k, \alpha)$-flattening configurations with $\bigcup_\ell K^{i\ell}_1 = N^i \cap C$.

Evidently $Q$ is closed, hence compact. By induction on $k$ there is a countable collection $\{((N^h_1, K^h_1), \ldots, (N^h_{k-1}, K^h_{k-1}))\}$ of $(k - 1, \alpha)$-flattening configurations such that $N^h_{k-1} \subset M$ for all $h$ and $\bigcup_h K^h_1 = Q$. Each $((N^h_1, K^h_1), \ldots, (N^h_{k-1}, K^h_{k-1}), (M, C))$ is a $(k, \alpha)$-flattening configuration, so we have the satisfactory countable collection

$$\{((N^{i\ell}_1, K^{i\ell}_1), \ldots, (N^{i\ell}_k, K^{i\ell}_k))\} \cup \{((N^h_1, K^h_1), \ldots, (N^h_{k-1}, K^h_{k-1}), (M, C))\}. \qquad \square$$

Note that $U$ itself is a full featured $C^{k,\alpha}$ manifold if we take $\mu_U = m$, $W_U = E$, $U_U = U$, $\psi_U$ the identity, $A_U = U$, $\rho_U$ the identity, and $Z_U = Z$. As we explained in Section 2, it suffices to prove Theorem 1 with the additional hypothesis that $C_{f,r} = \{\, c \in U : \ker Df(c) = Z \,\}$. Therefore the hypotheses of the following result are more general, so it completes the proof of Theorem 1 except for case c), which is Proposition 6 below.

**Proposition 5.** *If $M \subset E$ is a full featured $C^{k,\alpha}$ manifold, $f|_M$ is $C^{k,\alpha}$, $s \geq r$, and $S \subset C_{f,r} = \{\, c \in U : \ker Df(c) = T_c M \cap Z \,\}$ is $s$-sigmafinite, then $f(S)$ is $d(s)$-sigmafinite. If, in addition, either $S$ is $s$-null or $s > r$ and $M$ and $f$ are $C^{k,\alpha+}$, then $f(S)$ is $d(s)$-null.*

*Proof.* The special case $s = r = 0$ is easy: if $S$ is $s$-sigmafinite ($s$-null) then it is countable (empty) and $f(S)$ is countable (empty). Assume $s > 0$. Let $\{((N_1^j, K_1^j), \ldots, (N_k^j, K_k^j))\}$ be a $C^{k,\alpha}$ (or $C^{k,\alpha+}$ if $M$ and $f$ are $C^{k,\alpha+}$) Morse decomposition of $C_{f,r}$, as per Proposition 4. For each $j$ Proposition 3 implies that $\eta(N_1^j, K_1^j, f, k+\alpha) < \infty$ and $\eta(N_1^j, K_1^j, f, k+\alpha) = 0$ if $M$ and $f$ are $C^{k,\alpha+}$. If $\mu_{N_1^j} = r$, then Proposition 1 implies that $f(S \cap K_1^j)$ is $d(s)$-sigmafinite, and that it is $d(s)$-null if $s > r$. If $\mu_{N_1^j} > r$, then Proposition 2 implies that $f(S \cap K_1^j)$ is $d(s)$-finite, and that it is $d(s)$-null if $S$ is $s$-null or $s > r$ and $\eta(N_1^j, K_1^j, f, k+\alpha) = 0$. ☐

# 6 Open Domains

This section proves c) of Theorem 1. The proofs below are based on [2] and [34]. Now $f$ is $C^{k,\alpha}$, and our goal is:

**Proposition 6.** *$f(C_{f,r})$ is $d(m)$-null.*

Insofar as $C^k$ is the same thing as $C^{k,0+}$, the claim follows from Proposition 5 if $\alpha = 0$, so we assume that $\alpha > 0$. (The key consequence of this is that $k + \alpha > 1$.) Let $\{((N_1^j, K_1^j), \ldots, (N_k^j, K_k^j))\}$ be a $C^{k,\alpha}$ Morse decomposition of $C_{f,r}$, as per Proposition 4. It suffices to show, for a fixed $j$, that $f(K_1^j)$ is $d(m)$-null. If $\mu_{N_1^j} < m$, then $\mathcal{H}^m(K_1^j) = 0$, and Proposition 5 implies that $f(K_1^j)$ is $d(m)$-null. Therefore we may assume that $\mu_{N_1^j} = m$, and we can identify $N_1^j$ with $U$. To simplify notation we write $K$ in place of $K_1^j$.

Let $\tilde{K} \subset U$ be a compact neighborhood of $K$, and let $L$ be a Lipschitz constant for $f|_{\tilde{K}}$. Proposition 3 gives $\eta(U, K, f, k+\alpha) < \infty$. Choose $M > \eta(U, K, f, k+\alpha)$ and $\delta > 0$ such that $x \in \tilde{K}^j$ and $|f(x) - f(c)| \leq M|x - c|^{k+\alpha}$ whenever $c \in K$ and $x \in \tilde{K}$ with $x - c \in Z$ and $|x - c| \leq \delta$. Let $G = Lr^{1/2} + M(m - r)^{(k+\alpha)/2}$. Fix $\beta, \varepsilon > 0$.

Let $\mathcal{L}_m$ be Lebesgue outer measure on $E$.

**Lemma 6.** *For any $S \subset K$, $\mathcal{H}^{d(m)}(f(S)) \leq G^{d(m)} \mathcal{L}_m(S)$.*

*Proof.* The definition of $\mathcal{L}_m$ gives an open $V \subset U$ containing $S$ whose volume is less than $\mathcal{L}_m(S) + \varepsilon$. Consider a point $c_0 \in S$. Let $\lambda > 0$ be a number such that $P = \lambda^{1-k-\alpha}$ is an integer and $\lambda$ is small enough that $\sqrt{m-r}\lambda < \delta$, $G\lambda^{k+\alpha} < \beta$, and $\tilde{K} \cap V$ contains any cube of side length $\lambda$ centered at $c_0$. Let $C = A \times B$ be such a cube that is a product of a cube $B$ contained in $Z$ and a cube $A$ contained in its orthogonal complement. In the obvious way cover $A$ with $P^r$ closed cubes $A_h$ of sidelength $\lambda/P$. The diameter of $f((A_h \times B) \cap S)$ is bounded by $G\lambda^{k+\alpha}$ because if $c, c' \in (A_h \times B) \cap K$ with $c = (a, b)$ and $c' = (a', b')$, and $x = (a, b')$, then

$$|f(c') - f(c)| \leq |f(c') - f(x)| + |f(x) - f(c)| \leq L|a' - a| + M|b' - b|^{k+\alpha}$$

$$\leq Lr^{1/2}\lambda/P + M(m-r)^{(k+\alpha)/2}\lambda^{k+\alpha} = G\lambda^{k+\alpha}.$$

Summing over all $h$ such that $A_h \times B$ contains an element of $S$ gives

$$\mathcal{H}_\beta^{d(m)}(f(C \cap S)) \leq P^r (G\lambda^{k+\alpha})^{d(m)} = G^{d(m)} (\lambda^{1-k-\alpha})^r (\lambda^{k+\alpha})^{r + \frac{m-r}{k+\alpha}} = G^{d(m)} \mathcal{L}_m(C).$$

Since $c_0 \in S$ was arbitrary and $\lambda$ can be arbitrarily small, the Vitali covering theorem gives a countable disjoint collection $\{C_\ell\}$ of cubes $C_\ell \subset V$ satisfying this inequality with $\mathcal{L}_m(S \setminus \bigcup C_\ell) = 0$. Since $\mathcal{H}^m$ and $\mathcal{L}_m$ agree (up to normalization) on $E$, $f(S \setminus \bigcup C_\ell)$ is $d(m)$-null by Proposition 5. In general if $S_0 \cup S_1 = T$ and $\mathcal{H}^\tau(S_0) = 0$, then $\mathcal{H}^\tau_\beta(S_0) = 0$ for all $\beta > 0$, so $\mathcal{H}^\tau_\beta(T) \leq \mathcal{H}^\tau_\beta(S_0) + \mathcal{H}^\tau_\beta(S_1) = \mathcal{H}^\tau_\beta(S_1)$, but of course $\mathcal{H}^\tau_\beta(S_1) \leq \mathcal{H}^\tau_\beta(T)$. Since $\bigcup_\ell C_\ell \subset V$ we have

$$\mathcal{H}^{d(m)}_\beta(f(S)) = \mathcal{H}^{d(m)}_\beta(f(S \cap \bigcup C_\ell)) \leq G^{d(m)} \mathcal{L}_m(\bigcup C_\ell) < G^{d(m)}(\mathcal{L}_m(S) + \varepsilon). \qquad \square$$

Recall that $c \in E$ is a *density point* of $K$ if

$$\lim_{\ell \to \infty} \frac{\mathcal{L}_m(K \cap C_\ell)}{\mathcal{L}_m(C_\ell)} = 1$$

for any sequence $\{C_\ell\}$ of cubes containing $c$ and decreasing to $c$. Let $K_0$ be the set of density points of $K$ that are contained in $K$.

**Lemma 7.** *Any $c_0 \in K_0$ is contained in an arbitrarily small cube $C \subset U$ such that*

$$\mathcal{H}^{d(m)}_\beta(f(K \cap C)) < \varepsilon \mathcal{L}_m(C).$$

We also have $\mathcal{L}_m(C) < (1+\varepsilon)\mathcal{L}_m(K \cap C)$ if the cube $C$ given by this result is sufficiently small, because $c_0 \in K_0$. The Vitali theorem gives a countable disjoint collection $\{C_\ell\}$ of cubes satisfying these inequalities such that $\mathcal{L}_m(K_0 \setminus \bigcup C_\ell) = 0$. The Lebesgue density theorem ([16],[17, 2.9.11]) implies that $\mathcal{L}_m(K \setminus K_0) = 0$, so Lemma 6 implies that $\mathcal{H}^{d(m)}(f(K)) = \mathcal{H}^{d(m)}(f(K_0 \cap \bigcup C_\ell))$. We also have

$$\begin{aligned} \mathcal{H}^{d(m)}_\beta(f(K_0 \cap \bigcup C_\ell)) &\leq \sum_\ell \mathcal{H}^{d(m)}_\beta(f(K \cap C_\ell)) < \sum_\ell \varepsilon(1+\varepsilon)\mathcal{L}_m(K \cap C_\ell) \\ &= \varepsilon(1+\varepsilon)\mathcal{L}_m(K \cap \bigcup C_\ell) \leq \varepsilon(1+\varepsilon)\mathcal{L}_m(K). \end{aligned}$$

Since $\varepsilon$ and $\beta$ can be arbitrarily small, the proof of Proposition 6 is completed by:

*Proof of Lemma 7.* Let $J = 2^{k+\alpha} M(m-r)^{1+\frac{k+\alpha}{2}}$. Fix a positive integer $Q$ such that

$$G^{d(m)}/Q < \varepsilon/2 \quad \text{and} \quad 3^{d(m)} r^{r/2} L^r J^{\frac{m-r}{k+\alpha}} Q^{\frac{1-k-\alpha}{k+\alpha}(m-r)} < \varepsilon/2.$$

Choose $\lambda > 0$ such that $P = (L\sqrt{r}/J)(\lambda/Q)^{1-k-\alpha}$ is an integer and $\lambda$ is small enough that $\lambda < \delta Q/2\sqrt{m-r}$, $3JQ(\lambda/Q)^{k+\alpha} < \beta$, and $\mathcal{L}_m(K_j \cap C) \geq (1 - 1/Q^{m-r+1})\mathcal{L}_m(C)$ whenever $C$ is a cube of sidelength $\lambda$ containing $c_0$. (Of course $\lambda$ can be arbitrarily small.) Let $\gamma = JQ(\lambda/Q)^{k+\alpha}$, and note that $\gamma = L\sqrt{r}\lambda/P$.

Let $C = A \times B$ be a cube of sidelength $\lambda$ that contains $c_0$ where $B$ is a cube in $Z$ and $A$ is a cube in its orthogonal complement. Cover $A$ and $B$ with $P^r$ and $Q^{m-r}$ closed subcubes $A_g$ and $B_h$ of sidelengths $\lambda/P$ and $\lambda/Q$. Let

$$\mathcal{G} = \left\{\, g = 1, \ldots, P^r : \mathcal{L}_m(K \cap (A_g \times B)) > (1 - 1/Q^{m-r})\mathcal{L}_m(A_g \times B) \,\right\}.$$

Suppose that $g \in \mathcal{G}$ and $x = (a, b), x' = (a', b') \in A_g \times B$. We cover the line segment between $b$ and $b'$ with a minimal sequence of subcubes $B_{h_0}, \ldots, B_{h_N}$, with adjacent subcubes having nonempty intersection. Evidently $N \le (m-r)(Q-1) \le (m-r)Q-1$. By Fubini's theorem there is some $a^* \in A_g$ such that

$$\frac{\mathcal{L}_{m-r}(\{\, b \in B : (a^*, b) \in K \,\})}{\mathcal{L}_{m-r}(B)} > 1 - 1/Q^{m-r},$$

so that every $B_{h_\ell}$ contains some $b_\ell$ such that $(a^*, b) \in K$. Now $b$ and $b_0$ are in the same subcube, $b_{\ell-1}$ and $b_\ell$ are in intersecting subcubes for all $\ell \ge 1$, and $b_N$ and $b'$ are in the same subcube, so $|b - b_0|, |b_N - b'| \le \sqrt{m-r}\lambda/Q$ and $|b_{\ell-1} - b_\ell| \le 2\sqrt{m-r}\lambda/Q < \delta$ for $\ell \ge 1$. Setting $c_\ell = (a^*, b_\ell)$, we have

$$\begin{aligned}
|f(a^*, b) - f(a^*, b')| &\le |f(a^*, b) - f(c_0)| + \sum_{\ell=1}^{N} |f(c_{\ell-1}) - f(c_\ell)| + |f(c_N) - f(a^*, b')| \\
&\le M(|b - b_0|^{k+\alpha} + \sum_{\ell=1}^{N} |b_{\ell-1} - b_\ell|^{k+\alpha} + |b_N - b'|^{k+\alpha}) \\
&\le M(2^{-k-\alpha} + N + 2^{-k-\alpha})(2\sqrt{m-r}\lambda/Q)^{k+\alpha} \\
&\le 2^{k+\alpha} M(m-r)Q(\sqrt{m-r}\lambda/Q)^{k+\alpha} = JQ(\lambda/Q)^{k+\alpha} = \gamma.
\end{aligned}$$

We also have

$$|f(x) - f(a^*, b)|, |f(a^*, b') - f(x')| \le L \cdot \operatorname{diam} A_g = L\sqrt{r}\lambda/P = \gamma,$$

so $|f(x) - f(x')| \le 3\gamma < \beta$. Therefore

$$\begin{aligned}
\mathcal{H}^{d(m)}_{\beta}\Big(f\big(K \cap \bigcup_{g \in \mathcal{G}} A_g \times B\big)\Big) &\le P^r (3\gamma)^{d(m)} = 3^{d(m)} (\tfrac{L\sqrt{r}}{J} (\tfrac{\lambda}{Q})^{1-k-\alpha})^r (JQ(\tfrac{\lambda}{Q})^{k+\alpha})^{r+\frac{m-r}{k+\alpha}} \\
&= 3^{d(m)} r^{r/2} L^r J^{\frac{m-r}{k+\alpha}} Q^{\frac{1-k-\alpha}{k+\alpha}(m-r)} \lambda^m < \varepsilon \mathcal{L}_m(C)/2.
\end{aligned}$$

Since $\mathcal{L}_m(K_j \cap C) \ge (1 - 1/Q^{m-r+1})\mathcal{L}_m(C)$ we have $|\{\, g = 1, \ldots, P^r : g \notin \mathcal{G} \,\}|/P^r < 1/Q$. Of course $\mathcal{H}^{d(m)}_{\beta} \le \mathcal{H}^{d(m)}$, so Lemma 6 gives

$$\mathcal{H}^{d(m)}_{\beta}\Big(f\big(K \cap \bigcup_{g \notin \mathcal{G}} A_g \times B\big)\Big) \le G^{d(m)} \mathcal{L}_m\Big(K \cap \bigcup_{g \notin \mathcal{G}} A_g \times B\Big) \le G^{d(m)} \lambda^m / Q < \varepsilon \mathcal{L}_m(C)/2. \quad \square$$

# 7 Examples

To round out our picture of the subject, in this section we provide brief descriptions of the main methods that have been used to construct examples showing that weakenings of the hypotheses of Theorem 1 and related results, in various directions, are not possible. We will use one of them to construct examples showing that the differentiability hypotheses of Theorem 1 cannot be weakened.

We begin with Whitney's beautiful example [51], which can be thought of as the historical starting point of the topic. Figure 1 is a simplified version of a figure in [51], showing the construction of a path between the points $p$ and $q$, which are the midpoints of adjacent sides of a given square $Q$. We introduce four disjoint interior squares $Q_0, \ldots, Q_3$, where each $Q_i$ has an entry point $p_i$ and an exit point $q_i$, and there are five arcs $A_0, \ldots, A_4$, as shown. This process is repeated in each of $Q_0, \ldots, Q_3$, as shown in $Q_0$, and then iterated.

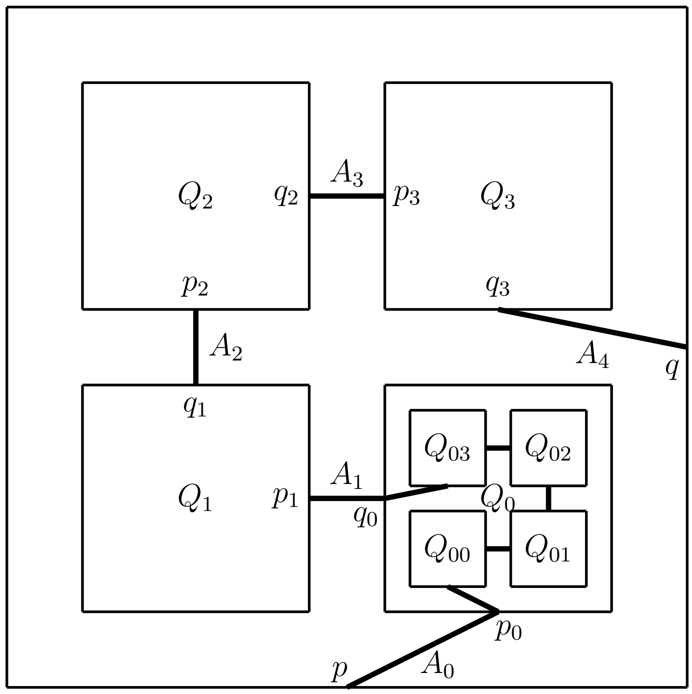


Figure 1

Formally, for $i_1, \ldots, i_k \in \{0, \ldots, 4\}$ let $\eta$ denote the multiindex $i_1 \cdots i_k$. If $\eta \in \{0, \ldots, 3\}^k$, then the process, as applied to $Q_\eta$, is as follows. We are given points of entry and exit $p_\eta$ and $q_\eta$, which are the midpoints of adjacent sides of $Q_\eta$. We introduce four smaller squares $Q_{\eta 0}, \ldots, Q_{\eta 3}$, and endow each $Q_{\eta i}$ with entry and exit points $p_{\eta i}$ and $q_{\eta i}$ that are the midpoints of adjacent sides of $Q_{\eta i}$, arranged so that the line seqments

$$A_{\eta 0} = \overline{p_\eta p_{\eta 0}}, \;\; A_{\eta 1} = \overline{q_{\eta 0} p_{\eta 1}}, \;\; A_{\eta 2} = \overline{q_{\eta 1} p_{\eta 2}}, \;\; A_{\eta 3} = \overline{q_{\eta 3}, p_{\eta 4}}, \;\; A_{\eta 4} = \overline{q_{\eta 4} q_\eta}$$

are as shown, and in particular do not intersect the interiors of the $Q_{\eta i}$ or each other. For $i_1, i_2, \ldots \in \{0, \ldots, 3\}$ let $\iota$ denote the infinite multiindex $i_1 i_2 \cdots$, and let $Q_\iota$ be the unique element of $\bigcap_k Q_{i_1 \cdots i_k}$.

Let $A$ be the union of all of the $A_m$ and the singletons $\{Q_\iota\}$; $A$ is the image of a continuous injective function $\gamma : [0, 1] \to Q$. (Giving an explicit parameterization by a suitable $\gamma$ is a tedious exercise, but not conceptually difficult.) Let $f : A \to \mathbb{R}$ be the function

$$f(x) = \begin{cases} \sum_{j=1}^k 4^{-j} i_j, & x \in A_{i_1 \cdots i_k}, \\ \sum_{j=1}^\infty 4^{-j} i_j, & x = Q_{i_1 i_2 \cdots}. \end{cases}$$

Note that $f$ is constant on each $A_\eta \cup A_{\eta 0}$ and each $A_{i_1 \cdots i_k} \cup A_{i_1 \cdots i_k - 1, 4}$. It follows easily that $f$ is continuous. Also, $f(\bigcup_\eta A_\eta)$ is countable and thus $\tau$-null for all $\tau > 0$.

We assume that for all $k$ and all $\eta \in \{0, \ldots, 3\}^k$, the ratio $\rho$ of the sidelength of $Q_\eta$ to the side lengths of $Q_{\eta 0}, \ldots, Q_{\eta 3}$ is the same as the ratio of the side length of $Q$ to the side lengths of $Q_0, \ldots, Q_3$. There are constants $C, D > 0$ such that for distinct $x, x' \in A'$, if $Q_{i_1 \cdots i_k}$ is the smallest cube that contains both $x$ and $x'$, then $|x - x'| \geq C\rho^{-k}$ and $|f(x) - f(x')| < D4^{-k}$. If $\rho < 4$, then the Whitney extension theorem [50] (or Theorem A1) implies that there is $C^1$ extension of $f$ to all of $\mathbb{R}^2$ such that each element of $A$ is a critical point, so the image of the set of critical points is all of $[0, 1]$. Furthermore we can compute the Hausdorff dimension of $A$; it is elegant to use the self similarity of $A$ and $A \cap Q_0$ to do this, but the Hausdorff dimension of $A$ is the Hausdorff dimension of $A' = \bigcap_k \bigcup_{\eta \in \{0,\ldots,3\}^k} Q_\eta$, which we analyze in more detail and greater generality below.

Whitney [51] points out that this construction extends easily to higher dimensions, which results in a higher order of differentiability, and the vanishing of all partial derivatives up to this order on the analogue of $A$. In addition one may take cartesian products of functions constructed in this way. Using these methods Norton [38, 39] showed that if $s > (k + \alpha)(n - r) + r$ (i.e., $d(s) > n$) and $m \geq (k + 1)(n - r) + r$, then there is a $C^{k+\alpha}$ function $g : \mathbb{R}^m \to \mathbb{R}^n$ and an $s$-null set $B \subset C_{g,r}$ such that $g(B)$ is not $n$-null.

Dubovitskiĭ [12, 13, 14] considers a variant of this construction that begins with the $p$-cube $[0, 1]^p$ and again passes to $2^p$ subcubes, $2^p$ subcubes of each of these subcubes, and so forth. Fix numbers $\ell > 1$ and $a > 0$ such that $\ell^{p-1} < 2$ and $pa < 1 - \frac{1}{\ell}$. In passing from the $k^{\text{th}}$ iteration to the next, we remove from each subcube the $p$ open "slices" of width $a/2^k\ell^k$ that have the center of the subcube in the middle and are parallel to the coordinate hyperplanes. Since the sidelength of the subcube is less than $2^{-k}$, the volume removed from the subcube is less than $p(2^{-k})^{p-1}\frac{a}{2^k\ell^k}$, and there are $2^{pk}$ subcubes after the $k^{\text{th}}$ iteration, so the total volume removed by this iteration is less than $pa/\ell^k$. Thus the $p$-dimensional measure of the set $L$ that remains after all iterations is not less than

$$1 - pa(1 + \ell^{-1} + \ell^{-2} + \cdots) = 1 - \frac{pa}{1 - \frac{1}{\ell}} > 0.$$

As in Whitney's example, $L$ is contained in the image $A = \gamma([0, 1])$ of an injective curve $\gamma : [0, 1] \to [0, 1]^p$ that, within each subcube, proceeds through the subcubes of the next level in order. A function on $A$ can be defined by setting $\Phi(\gamma(t)) = \mathcal{H}^p(\gamma([0, t]))$. If $x$ and $x'$ are points in $A$ that are in the same cube after $k$ iterations and different cubes after $k + 1$ stages, then $|x - x'| > a/2^k\ell^k$, and

$$|\Phi(x) - \Phi(x')| < \frac{1}{2^{pk}} = \frac{1}{2^{pk}}\Big(\frac{2^k\ell^k}{a}\Big)^{p-1}\Big(\frac{a}{2^k\ell^k}\Big)^{p-1} < \frac{1}{a^{p-1}}(\ell^{p-1}/2)^k|x - x'|^{p-1}.$$

Now Whitney's extension theorem gives a $C^{p-1}$ extension of $\Phi$ to $\mathbb{R}^p$ whose partial derivatives up to order $p - 1$ vanish at each point of $A$. Dubovitskiĭ attributed this example to Dmitrii Menshov. Let $\psi_p : [0, 1]^p \to [0, 1]$ be $\mathcal{H}^p(A)^{-1}$ times the extended $\Phi$.

Dubovitskiĭ constructed the following example to show that the minimum degree of differentiability in the result described in the introduction is best possible. Let $m$, $n$, and

$\nu$ be arbitrary natural numbers such that $m - n - \nu > 0$, and let $k = m - n - \nu$. Let $f : [0,1]^m \to [0,1]^n$ be given by setting

$$f_1(x_1, \dots, x_m) = \psi_{k+1}(x_1, \dots, x_{k+1})$$

and $f_i(x_1, \dots, x_m) = x_{m-n+i}$ for $i = 2, \dots, n$. Of course $f$ is $C^k$. For any $u \in [0,1]^n$ there is a point $(\overline{x}_1, \dots, \overline{x}_{k+1}) \in \psi_{k+1}^{-1}(u_1)$ such that

$$D\psi_{k+1}(\overline{x}_1, \dots, \overline{x}_{k+1}) = 0.$$

Of course

$$(\overline{x}_1, \dots, \overline{x}_{k+1}, x_{k+2}, \dots, x_{m-n+1}, u_2, \dots, u_n) \in f^{-1}(u)$$

for all $x_{k+2}, \dots, x_{m-n+1} \in [0,1]$. Since the projection onto these coordinates is Lipschitz, it follows that $\mathcal{H}^\nu(C_f \cap f^{-1}(u)) > 0$. In particular,

$$\mathcal{H}^n(\{\, u \in [0,1]^n : \mathcal{H}^\nu(C_f \cap f^{-1}(u)) > 0 \,\}) > 0.$$

We now describe and apply the style of example that uses generalized Cantor sets and maps between them. This approach seems to have been originated by Federer [17, 3.4.3], and has been used in [4, 34, 35]. (Grinberg [23] provides an extremely simple instance of this construction.) It is closely related to Whitney's method, but achieves somewhat greater simplicity and generality by discarding inessential features.

For $0 < \varphi < 1$ the Cantor set $K_\varphi$ is $\bigcap_{j=0}^\infty C_\varphi(j)$ where $C_\varphi(0) = [0,1]$ and, for each $j \geq 1$, $C_\varphi(j)$ is obtained from $C_\varphi(j-1)$ by removing, from each of the $2^{j-1}$ closed intervals $I$ constituting $C_\varphi(j-1)$, the open middle interval whose length is $\varphi$ times the length of $I$. Fix a second number $\psi \in (0,1)$. For each $j$ let the intervals that constitute $[0,1] \setminus C_\varphi(j)$ be $U_1(j), \dots, U_{2^j-1}(j)$, and let $V_1(j), \dots, V_{2^j-1}(j)$ be the corresponding intervals constituting $[0,1] \setminus C_\psi(j)$. Let $g : K_\varphi \to K_\psi$ be the continuous extension to all of $K_\varphi$ of the function that maps the endpoints of each $U_h(j)$ to the corresponding endpoints of $V_h(j)$. Evidently $g$ is a homeomorphism.

Fix an integer $k \geq 1$ and $\alpha \in [0,1]$. If $u, v \in K_\varphi$ and $j$ is the first integer such that $u$ and $v$ are in different intervals in $C_\varphi(j+1)$, then $|u - v| \geq \varphi(\frac{1-\varphi}{2})^j$ and $|g(u) - g(v)| \leq (\frac{1-\psi}{2})^j$, so that

$$\frac{|g(u) - g(v)|}{|u - v|^{k+\alpha}} \leq \frac{1}{\varphi^{k+\alpha}} \Big( (\frac{1-\psi}{2}) \Big/ (\frac{1-\varphi}{2})^{k+\alpha} \Big)^j.$$

Suppose that $(\frac{1-\varphi}{2})^{k+\alpha} \geq \frac{1-\psi}{2}$. There is $M = \varphi^{-k-\alpha} > 0$ such that $|g(u) - g(v)| \leq M|u - v|^{k+\alpha}$ for all $u, v \in K_\varphi$, and if $(\frac{1-\varphi}{2})^{k+\alpha} > \frac{1-\psi}{2}$, then for any $\varepsilon > 0$ there is $\delta > 0$ such that $|g(u) - g(v)| < \varepsilon |u - v|^{k+\alpha}$ whenever $u, v \in K_\varphi$ with $|u - v| < \delta$.

Fix integers $m$, $n$, and $r$ with $1 \leq r < n \leq m$, and let

$$S = [0,1]^r \times K_\varphi^{n-r} \times \{0\} \subset \mathbb{R}^m \quad \text{and} \quad T = [0,1]^r \times K_\psi^{n-r} \subset \mathbb{R}^n.$$

(Here 0 is the origin of $\mathbb{R}^{m-n}$.) Let $f : S \to \mathbb{R}^n$ be the map

$$f(a) = (a_1, \dots, a_r, g(a_{r+1}), \dots, g(a_n)).$$

Since $g$ is a homeomorphism, so is $f$. For each $a \in S$ let $P_a : \mathbb{R}^m \to \mathbb{R}^m$ be the linear polynomial

$$P_a(x) = (x_1, \ldots, x_r, g(a_{r+1}), \ldots, g(a_m)).$$

Then

$$|P_a(b) - P_b(b)| = \max_{s=r+1,\ldots,n} |g(a_s) - g(b_s)| \leq M|a-b|^{k+\alpha}$$

for all $a, b \in S$, and if $(\frac{1-\varphi}{2})^{k+\alpha} > \frac{1-\psi}{2}$, then for any $\varepsilon > 0$ there is $\delta > 0$ such that $|P_a(b) - P_b(b)| < \varepsilon|a-b|^{k+\alpha}$ whenever $a, b \in S$ with $|a-b| < \delta$. For all $a, b \in S$, $P_a(\cdot)$ and $P_b(\cdot)$ differ by a constant, so, for all $i = 1, \ldots, k$, $D^i P_a(b) = D^i P_b(b)$. If $\alpha > 0$, then Theorem A1 implies that there is a $C^{k,\alpha}$ function $\tilde{f} : \mathbb{R}^m \to \mathbb{R}^n$ such that $D^i \tilde{f}(a) = D^i P_a(a)$ for all $i = 0, \ldots, k$ and all $a \in S$. If $(\frac{1-\varphi}{2})^{k+\alpha} > \frac{1-\psi}{2}$, then Theorem A1 implies that there is a $C^{k,\alpha+}$ function $\tilde{f} : \mathbb{R}^m \to \mathbb{R}^n$ such that $D^i \tilde{f}(a) = D^i P_a(a)$ for all $i = 0, \ldots, k$ and all $a \in S$. In either case $P_a(a) = f(a)$ for $a \in S$, so $\tilde{f}$ is an extension of $f$, and $D^i \tilde{f}(a) = 0$ for all $i = 1, \ldots, k$ and all $a \in S$. In particular $S \subset C_{\tilde{f},r}$.

We now compute the Hausdorff dimensions of $S$ and $T$. Let $\sigma = -\ln 2/\ln(\frac{1-\varphi}{2})$. We have $2(\frac{1-\varphi}{2})^\sigma = 1$, and thus $2^j(\frac{1-\varphi}{2})^{j\sigma} = 1$ for all $j$, so the sum, over the intervals constituting $C_\varphi(j)$, of the lengths raised to the power $\sigma$ is 1. Differentiation of the expression $(\delta-\gamma)^\sigma - (\frac{1-\varphi}{2} - \gamma)^\sigma - (\delta - \frac{1+\varphi}{2})^\sigma$ with respect to $\gamma$ and $\delta$ shows that it is positive when $\gamma \in [0, \frac{1-\varphi}{2}]$, $\delta \in [\frac{1+\varphi}{2}, 1]$, and either $\gamma > 0$ or $\delta < 1$. From this it follows easily that for any cover of $K_\varphi$ by finitely many intervals, the sum of the lengths of intervals raised to the power $\sigma$ is decreased by refining it to some $C_\varphi(j)$. We conclude that $\mathcal{H}^\sigma(K_\varphi) = 1$. Let $\tau = -\ln 2/\ln(\frac{1-\psi}{2})$. The same argument gives $\mathcal{H}^\tau(K_\psi) = 1$.

In general, if $X$ and $Y$ are metric spaces, $X \times Y$ is endowed with a metric such that the the projections $X \times Y \to X$ and $X \times Y \to Y$ are 1-Lipschitz, and $\mathcal{H}^\kappa(X), \mathcal{H}^\lambda(Y) < \infty$, then $\mathcal{H}^{\kappa+\lambda}(X \times Y) \geq \mathcal{H}^\kappa(X) \cdot \mathcal{H}^\lambda(Y)$ ([17, 2.10.27]). This inequality may hold strictly (examples are given in [17, 2.10.29]) but insofar as, for $\ell \geq 2$, $C_\varphi(j)^\ell$ is covered by $2^{\ell j}$ cubes of diameter $\sqrt{\ell}(\frac{1-\varphi}{2})^j$, and $2^{\ell j}(\sqrt{\ell}(\frac{1-\varphi}{2})^j)^{\ell\kappa} = \ell^{\ell\kappa/2}(2^j(\frac{1-\varphi}{2})^j)^{\ell\kappa} = \ell^{\ell\kappa/2}$, we find that $1 \leq \mathcal{H}^{\ell\kappa}(K_\varphi^\ell) \leq \ell^{\ell\kappa/2}$. A straightforward extension of this argument shows that $1 \leq \mathcal{H}^{h+\ell\kappa}([0,1]^h \times K_\varphi^\ell) < \infty$ for all nonnegative integers $h$ and $\ell$. Thus $1 \leq \mathcal{H}^{r+(n-r)\sigma}(S) < \infty$ and $1 \leq \mathcal{H}^{r+(n-r)\tau}(T) < \infty$.

Suppose that $(\frac{1-\varphi}{2})^{k+\alpha} > \frac{1-\psi}{2}$. Then $(k+\alpha)\ln(\frac{1-\varphi}{2}) = \ln(\frac{1-\psi}{2})$, so if $s = r + (n-r)\sigma$, then $d(s) = r + (s-r)/(k+\alpha) = r + (n-r)\tau$. Since $\sigma$ may be any positive number, $\tilde{f}$ is an example of how, in Theorem 1, for any $s > r$ and any range dimension $n = r+1, \ldots, m$, $f(S)$ may fail to be $d(s)$-null if none of a), b), and c) hold.

Now suppose that $(\frac{1-\varphi}{2})^{k+\alpha} > \frac{1-\psi}{2}$. Now $d(s) < r + (n-r)\tau$, so $\tilde{f}(S)$ is not $d(s)$-sigmafinite. By varying $\varphi$ and $\psi$ we can obtain a complete set of examples showing that, for any $n = r+1, \ldots, m$, $f(S)$ may fail to be $t$-sigmafinite whenever $U \subset E$ is open, $f : E \to Y$ is $C^{k,\alpha+}$, $S \subset C_{f,r}$ is $s$-sigmafinite for some $s > r$ and $t < d(s)$. (Of course this construction and the one in the previous paragraph also give satisfactory examples when $m < \dim Y$.)

In either case if, for $y \in \mathbb{R}^r$, we let $h_y : \mathbb{R}^{n-r} \to \mathbb{R}^{n-r}$ be the function $h_y(z) = (\tilde{f}_{r+1}(y, z, 0), \ldots, \tilde{f}_n(y, z, 0))$, then $D^i h_y(z) = 0$ for all $i = 1, \ldots, k$ and all $z \in K_\varphi^{m-r}$.

Thus one cannot obtain stronger bounds on the Hausdorff dimension of $\tilde{f}(S)$ by requiring the derivatives of $\tilde{f}$ up to order $k$ to vanish in the null spaces of the derivative of $\tilde{f}$.

# Appendix A: $C^{k,\alpha}$ and $C^{k,\alpha+}$ Whitney Extensions

In this appendix we provide a version of the Whitney extension theorem for $C^{k,\alpha}$ and $C^{k,\alpha+}$ functions. This result has also been proved by [38], and a version of it appears in [49]. Our argument, which is adapted from [17, 3.1.14], is relatively succinct.

**Theorem A1.** *Suppose that $k$ is a nonnegative integer, $\alpha \in [0,1]$, $A$ is a closed subset of $E$, $U = E \setminus A$, and, for each $a \in A$, $P_a : E \to Y$ is a polynomial function of degree $k$. For each nonempty compact $K \subset A$ and $\delta > 0$ let*

$$\varepsilon_K(\delta) = \sup_{i=0,\dots,k,\, a,b\in K,\, a\neq b,\, |a-b|\leq\delta} \frac{|D^i P_a(b) - D^i P_b(b)|}{|a-b|^{k+\alpha-i}},$$

*and let $M_K = \varepsilon_K(1)$. If $\alpha > 0$ and $M_K < \infty$ for all $K$, then there is a $C^{k,\alpha}$ function $f : E \to Y$ such that $D^i f(a) = D^i P_a(a)$ for all $a \in A$ and all $i = 0, \dots, k$. If $\lim_{\delta\to 0} \varepsilon_K(\delta) = 0$ for all $K$, then there is a $C^{k,\alpha+}$ function $f : E \to Y$ such that $D^i f(a) = D^i P_a(a)$ for all $a \in A$ and all $i = 0, \dots, k$. In either case $f|_U$ is $C^\infty$.*

*Proof.* For $x \in E$ let $h(x) = \min_{y\in A} |x - y|$ be the distance from $x$ to $A$, and for $\delta > 0$ let $\mathbf{U}_\delta(x) = \{\, y \in E : |x - y| < \delta \,\}$ be the open $\delta$-ball centered at $x$.

Zorn's lemma implies the existence of a maximal $S \subset U$ such that the balls $\mathbf{U}_{h(s)/10}(s)$ for $s \in S$ are disjoint. We claim that $\bigcup_{s\in S} \mathbf{U}_{2h(s)/5}(s) = U$. Aiming at a contradiction, suppose $s^* \in U \setminus \bigcup_{s\in S} \mathbf{U}_{2h(s)/5}(s)$. Maximality implies that $\mathbf{U}_{h(s^*)/10}(s^*)$ intersects some $\mathbf{U}_{h(s)/10}(s)$, in which case $h(s^*) < h(s^*)/10 + h(s)/10 + h(s)$, but $s^* \notin \mathbf{U}_{2h(s)/5}(s)$ and $\mathbf{U}_{h(s^*)/10}(s^*) \cap \mathbf{U}_{h(s)/10}(s) \neq \emptyset$ imply that $h(s^*)/10 > 3h(s)/10$, which is impossible.

Let $\mu : \mathbb{R} \to [0,1]$ be a $C^\infty$ function such that $\mu^{-1}((0,1]) = (-1,1)$ and $\mu^{-1}(1) = [-1/2, 1/2]$. Let $u_0 : E \to [0,1]$ be the function $u_0(x) = \mu(|x|)$. This function is $C^\infty$ on $E \setminus \{0\}$ and constant on a neighborhood of 0, so it is $C^\infty$. For $s \in S$ let $u_s : E \to [0,1]$ be the function $u_s(x) = u_0(2(x-s)/h(s))$, which of course is $C^\infty$, positive on $\mathbf{U}_{h(s)/2}(s)$, and zero elsewhere. The open cover $\{\, \mathbf{U}_{h(s)/2}(s) : s \in S \,\}$ of $U$ is locally finite, so the function $\sigma(x) = \sum_s u_s(x)$ is well defined on $U$, $C^\infty$, and positive everywhere. For each $s$ let $v_s : U \to [0,1]$ be the function $v_s(x) = u_s(x)/\sigma(x)$. Of course $v_s$ is $C^\infty$, it is positive on $\mathbf{U}_{h(s)/2}(s)$ and vanishes elsewhere, and $\sum_s v_s \equiv 1$, which is to say that $\{\, v_s : s \in S \,\}$ is a $C^\infty$ *partition of unity* for $U$ subordinate to the cover $\{\, \mathbf{U}_{h(s)/2}(s) : s \in S \,\}$.

For each $i = 0, \dots, k$ let $D_i = \max_x |D^i u_0(x)|$. Clearly $|D^i u_s(x)| \leq D_i h(s)^{-i}$ for all $s$, $i$, and $x$. Let $B = \min_{|x|\leq 4/5} u_0(x)$. Since the balls $\mathbf{U}_{2h(s)/5}(s)$ cover $U$ we have $\sigma(x) \geq B$ for all $x$. If $s, s' \in S$ and $\mathbf{U}_{h(s')/2}(s') \cap \mathbf{U}_{h(s)/2}(s) \neq \emptyset$, then $h(s) \leq 3h(s')/2 + h(s)/2$, so $h(s') \geq h(s)/3$. Since the balls $\mathbf{U}_{h(s)/10}(s)$ are disjoint, it follows that there is a constant $L$, depending only on the dimension $m$ of $E$, such that for each $s$,

$$|\{\, s' \in S : \mathbf{U}_{h(s')/2}(s') \cap \mathbf{U}_{h(s)/2}(s) \neq \emptyset \,\}| \leq L.$$

The standard formulas for differentiation, applied to the quotient $v_s = u_s / \sum_{s'} u_{s'}$, and the fact that $h(x) > h(s)/2$ if $x \in \mathbf{U}_{h(s)/2}(s)$, imply that for each $i = 0, \ldots, k$ there is a constant $V_i$, depending only on $m$, $B$, and $D_0, \ldots, D_i$, such that $|D^i v_s(x)| \le V_i h(x)^{-i}$ for all $s$ and $x \in U$.

For each $s$ choose $\xi(s) \in A$ such that $|\xi(s) - s| = h(s)$. Let $f : E \to Y$ be the function given by

$$f(x) = \begin{cases} P_x(x), & x \in A, \\ \sum_s v_s(x) P_{\xi(s)}(x), & x \in U. \end{cases}$$

Clearly $f|_U$ is $C^\infty$. It remains to show that it has the desired behavior on and near $A$. Fix an $a \in A$. Let $K \subset A$ be a compact set that contains $a$ in its relative interior, let $\delta > 0$ be small enough that $A \cap \mathbf{U}_{7\delta}(a) \subset K$. Fix a point $x \in \mathbf{U}_\delta(a)$, and let $b \in A$ be a point such that $|x - b| = h(x)$. Of course $b \in \mathbf{U}_{2\delta}(a)$, so $b \in K$.

Suppose that $s \in S$ and $x \in \mathbf{U}_{h(s)/2}(s)$. In order to bound $|\xi(s) - b|$ we observe that

$$|s - x| \le \tfrac{1}{2} h(s) \le \tfrac{1}{2} |s - b| \le \tfrac{1}{2}(|s - x| + |x - b|),$$

so $|s - x| < |x - b|$, from which we obtain

$$|s - \xi(s)| \le |s - b| \le |s - x| + |x - b| \le 2|x - b|$$

and

$$|\xi(s) - b| \le |\xi(s) - s| + |s - x| + |x - b| \le 4|x - b|.$$

In addition, $|\xi(s) - a| \le 6|x - a|$, and thus $\xi(s) \in K$.

For the time being we work with a fixed $i = 0, \ldots, k$. For any $c \in K$ Taylor's theorem gives $D^i P_c(x) = \sum_{j=0}^{k-i} \frac{1}{j!} D^{i+j} P_c(b)(x - b, \ldots, x - b)$ and thus

$$\begin{aligned} |D^i P_c(x) - D^i P_b(x)| &\le \sum_{j=0}^{k-i} \tfrac{1}{j!} |x - b|^j \cdot |D^{i+j} P_c(b) - D^{i+j} P_b(b)| \\ &\le (k - i + 1) \max\{|x - b|, |c - b|\}^{k+\alpha-i} \varepsilon_K(|c - b|). \end{aligned}$$

Note that $|b - a| \le |b - x| + |x - a| \le 2|x - a|$, so when $c = a$ we have

$$|D^i P_a(x) - D^i P_b(x)| \le (k - i + 1)(2|x - a|)^{k+\alpha-i} \varepsilon_K(|a - b|).$$

If $\kappa : E^i \to \mathbb{R}$ and $\lambda : E^j \to Y$ are multilinear and symmetric, let $\kappa \odot \lambda : E^{i+j} \to V$ be the symmetric multilinear function

$$\kappa \odot \lambda(v) = \sum \kappa(v_h) \cdot \lambda(v_{-h})$$

where the sum is over all $1 \le h_1 < \cdots < h_i \le i + j$, $v_h = (v_{h_1}, \ldots, v_{h_i})$, and $v_{-h} = (v_1, \ldots, v_{h_1 - 1}, v_{h_1 + 1}, \ldots, v_{h_i - 1}, v_{h_i + 1}, \ldots, v_{i+j})$. Note that $|\kappa \odot \lambda| \le \binom{i+j}{i} |\kappa||\lambda|$. This notation allows us to write

$$D^i f(x) = \sum_s \sum_{j=0}^{i} D^{i-j} v_s(x) \odot D^j P_{\xi(s)}(x)$$

for $x \in U$ and $i = 0, \ldots, k$. We now have

$$|D^i f(x) - D^i P_b(x)| = \Big|\sum_s \sum_{j=0}^{i} D^{i-j} v_s(x) \odot [D^j P_{\xi(s)}(x) - D^j P_b(x)]\Big|$$

$$\leq \sum_s \sum_{j=0}^{i} \binom{i+j}{i} |D^{i-j} v_s(x)| \cdot |D^j P_{\xi(s)}(x) - D^j P_b(x)|$$

$$\leq \sum_s \sum_{j=0}^{i} \binom{i+j}{i} V_{i-j} h(x)^{j-i} \cdot (k-j+1) \max\{|x-b|, |\xi(s)-b|\}^{k+\alpha-j} \varepsilon_K(|\xi(s)-b|)$$

$$\leq \sum_s \sum_{j=0}^{i} \binom{i+j}{i} V_{i-j} (k-j+1) 4^{k+\alpha-j} |x-b|^{k+\alpha-i} \varepsilon_K(4|x-b|).$$

Thus there is a constant $N_i$, depending on $m$, $k$, $i$, and $V_0, \ldots, V_i$, such that

$$|D^i f(x) - D^i P_b(x)| \leq N_i |x-b|^{k+\alpha-i} \varepsilon_K(4|x-b|).$$

Combining this and the result above,

$$\begin{aligned} |D^i f(x) - D^i P_a(x)| &\leq |D^i P_a(x) - D^i P_b(x)| + |D^i f(x) - D^i P_b(x)| \\ &\leq ((k-i+1)2^{k+\alpha-i} + N_i)|x-a|^{k+\alpha-i} \varepsilon_K(4|x-a|). \end{aligned}$$

We now show that $D^i f(a) = D^i P_a(a)$ for all $i = 0, \ldots, k$. This holds by construction when $i = 0$. Arguing by induction, suppose that $0 \leq i < k$ and it has already been established that $D^i f(a) = D^i P_a(a)$. We have

$$|D^i f(x) - D^i f(a) - D^{i+1} P_a(a)(x-a)| \leq |D^i f(x) - D^i P_a(x)|$$

$$+ |D^i P_a(x) - D^i P_a(a) - D^{i+1} P_a(a)(x-a)| = o(|x-a|)$$

because the inequality above gives $|D^i f(x) - D^i P_a(x)| = o(|x-a|)$ and $P_a$ is a polynomial.

Finally, according to whether $\alpha > 0$ and $M_K < \infty$ for all $K$, or $\lim_{\delta \to 0} \varepsilon_K(\delta) = 0$ for all $K$, the inequality above implies that $D^k f$ is $C^{0,\alpha}$ or $C^{0,\alpha+}$, so $f$ is $C^{k,\alpha}$ or $C^{k,\alpha+}$. □

# Appendix B: Properties of $C^{k,\alpha}$ and $C^{k,\alpha+}$ Functions

This Appendix establishes some basic facts about $C^{k,\alpha}$ and $C^{k,\alpha+}$ functions, slightly expanding the content of [38, Appendix]. In order to minimize space and complexity the results are stated and proved only for $C^{k,\alpha}$ functions, but they are valid with $C^{k,\alpha+}$ replacing $C^{k,\alpha}$ everywhere, so they should be understood as pertaining equally to that differentiability class. To begin with we note that if $g$ and $h$ are composable $C^1$ functions, then

$$D(h \circ g)(x') - D(h \circ g)(x) = (Dh(g(x')) - Dh(g(x))) \circ Dg(x') + Dh(g(x)) \circ (Dg(x') - Dg(x)).$$

**Lemma B1.** *The composition of a $C^k$ map $g$ with a $C^{k-1,\alpha}$ map $h$ is $C^{k-1,\alpha}$.*

*Proof.* The case $k = 1$ is easy. Suppose the result has already been established with $k-1$ in place of $k$, so $Dh \circ g$ is $C^{k-2,\alpha}$. Of course $Dg$ is $C^{k-1}$, so, in view of the calculation above, $D(h \circ g)$ is $C^{k-2,\alpha}$, and we are done. □

**Lemma B2.** *The composition of a $C^{k,\alpha}$ map $g$ with a $C^{k,\alpha}$ map $h$ is $C^{k,\alpha}$.*

*Proof.* The case $k = 1$ follows directly from the calculation above. Suppose the result has already been established with $k-1$ in place of $k$. The last result implies that $Dh \circ g$ is $C^{k-1,\alpha}$, and of course $Dg$ is $C^{k-1,\alpha}$, so, by the calculation above, $D(h \circ g)$ is $C^{k-1,\alpha}$. □

**Lemma B3.** *If $U \subset \mathbb{R}^m$ is open, $f\colon U \to \mathbb{R}^m$ is $C^{k,\alpha}$, and $x_0 \in U$ is not a critical point of $f$, then there is an open $V \subset U$ containing $x_0$ such that $f|_V$ is injective and $(f|_V)^{-1}$ is $C^{k,\alpha}$.*

*Proof.* The standard inverse function theorem gives an open $V \subset U$ containing $x_0$ such that $f|_V$ is injective and $(f|_V)^{-1}$ is $C^k$. The chain rule gives $D(f \circ f^{-1}) = (Df \circ f^{-1}) \circ Df^{-1}$, so $Df^{-1} = \mathrm{Inv} \circ Df \circ f^{-1}$, which is $C^{k-1,\alpha}$ by Lemma B1. □

**Lemma B4.** *If $U \subset \mathbb{R}^m \times \mathbb{R}$ is open, $f\colon U \to \mathbb{R}$ is $C^{k,\alpha}$, $(x_0, y_0) \in f^{-1}(0)$, and $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, then there is an open $V \subset \mathbb{R}^m$ containing $x_0$ and a $C^{k,\alpha}$ function $g\colon V \to \mathbb{R}$ such that $\{\, (x, g(x)) : x \in V \,\}$ is a neighborhood of $(x_0, y_0)$ in $f^{-1}(0)$.*

*Proof.* The standard implicit function theorem gives $V$ and a $C^k$ function $g$ satisfying the asserted condition. A simple computation shows that $Dg(x) = -\frac{\partial f}{\partial x}(x, g(x))/\frac{\partial f}{\partial y}(x, g(x))$ is (by Lemma B2) a quotient of $C^{k-1,\alpha}$ functions, and is consequently $C^{k-1,\alpha}$. □

# Acknowledgements

I am grateful for helpful correspondence with Mikhail Korobkov, Frank Morgan, and Ravi Vakil. Suggestions by an anonymous referee have significantly improved the paper.

# References

[1] G. Alberti, Generalized $N$-property and Sard theorem for Sobolev maps, *Rend. Lincei Mat. Appl.* **23** (2012), 477–491.

[2] S. M. Bates, Toward a precise smoothness hypothesis in Sard's theorem, *Proc. Amer. Math. Soc.* **117** (1993), 279–283.

[3] S. M. Bates and C. G. T. d. A. Moreira, *A Generalization of Sard's Theorem*, Preprint, 2000.

[4] S. M. Bates and C. G. T. d. A. Moreira, De nouvelles perspectives sur le théorème de Morse-Sard, *C. R. Acad. Sci. Paris - Series I* **332** (2001), 13–17.

[5] B. Bojarski, P. Hajłasz, and P. Strzelecki, Sard's theorem for mappings in Hölder and Sobolev spaces, *Manuscripta Math.* **118** (2005), 383–397.

[6] J. Bourgain, M. V. Korobkov, and J. Kristensen, On the Morse-Sard property and level sets of Sobolev and BV functions, *Rev. Mat. Iberoam* **29** (2013), 1–23.

[7] J. Bourgain, M. V. Korobkov, and J. Kristensen, On the Morse-Sard property and level sets of $W^{n,1}$ Sobolev functions on $\mathbb{R}^n$, *J. Reine Angew. Math.* **700** (2015), 93–112.

[8] A. Brown, Functional dependence, *Trans. Amer. Math. Soc.* **38** (1935), 379–394.

[9] A. Constantin, Critical values of differentiable maps, *Expositiones Mathematicae* **14** (1996), 93–95.

[10] L. De Pascale, The Morse-Sard theorem in Sobolev spaces, *Indiana Univ. Math. J.* **50** (2001), 1371–1386.

[11] G. Debreu, Economies with a finite set of equilibria, *Econometrica* **38** (1970), 387–392.

[12] A. Y. Dubovitskiĭ, On the structure of level sets of differentiable mappings of an $n$-dimensional cube into a $k$-dimensional cube, *Izv. Akad. Nauk SSSR Ser. Mat.* **21** (1957), 371–408.

[13] A. Y. Dubovitskiĭ, On the point sets of a given degenerate Jacobian matrix of infinitely differentiable mappings, *Izv. Akad. Nauk SSSR Ser. Mat.* **26** (1962), 489–494.

[14] A. Y. Dubovitskiĭ, On the set of degenerate points, *Izv. Akad. Nauk SSSR Ser. Mat.* **31** (1967), 27–36.

[15] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, Taylor and Francis Group, Boca Raton, 2015.

[16] C.-A. Faure, A short proof of Lebesgue's density theorem, *Amer. Math. Monthly* **109** (2002), 194–196.

[17] H. Federer, *Geometric Measure Theory*, Springer, New York, 1969.

[18] H. Federer, Two theorems in geometric measure theory, *Bull. Amer. Math. Soc.* **72** (1966), 719.

[19] A. Ferone, M. V. Korobkov, and A. Roviello, Morse-Sard theorem and Luzin $N$-property: a new synthesis result for Sobolev spaces, *Siberian Math. J.* **60** (2019), 916–926.

[20] A. Ferone, M. V. Korobkov, and A. Roviello, On the Luzin $N$-property and the uncertainty principle for Sobolev mappings, *Anal. PDE* **12** (2019), 1149–1175.

[21] A. Ferone, M. V. Korobkov, and A. Roviello, On some universal Morse-Sard type theorems, *J. Math. Pures Appl.* **139** (2020), 1–34.

[22] A. Figalli, A simple proof of the Morse-Sard theorem in Sobolev spaces, *Proc. Amer. Math. Soc.* **136** (2008), 3675–3681.

[23] E. L. Grinberg, On the smoothness hypothesis in Sard's theorem, *Amer. Math. Monthly* **92** no. 10 (1985), 733–734.

[24] P. Hajłasz, M. V. Korobkov, and J. Kristensen, A bridge between Dubovitskiĭ-Federer theorems and the coarea formula, *J. Funct. Anal.* **272** (2017), 1265–1295.

[25] P. Hajłasz and S. Zimmerman, The Dubovitskiĭ-Sard theorem in Sobolev spaces, *Indiana Univ. Math. J.* **66** (2017), 705–723.

[26] J. C. Harsanyi, Oddness of the number of equilibrium points: A new proof, *Int. J. Game Theory* **2** (1973), 235–250.

[27] R. Kaufman, Representation of linear sets as critical sets, *Proc. Amer. Math. Soc* **25** (1970), 884–889.

[28] M. V. Korobkov and J. Kristensen, On the Morse-Sard theorem for the sharp case of Sobolev mappings, *Indiana Univ. Math. J.* **63** (2014), 1703–1724.

[29] M. V. Korobkov and J. Kristensen, The trace theorem, the Luzin N- and Morse-Sard properties for the sharp case of Sobolev-Lorentz mappings, *J. Geom. Anal.* **28** (2018), 2834–2856.

[30] D. Kreps and R. Wilson, Sequential equilibrium, *Econometrica* **50** (1982), 863–894.

[31] I. Kupka, Counterexample to the Morse-Sard theorem in the case of infinite dimensional manifolds, *Proc. Amer. Math. Soc.* **16** (1965), 954–957.

[32] M. Kučera, Hausdorff measures of the set of critical values of functions of the class $c^{k,\lambda}$, *Comment. Math. Univ.Carolin.* **13** (1972), 333–350.

[33] J. Milnor, *Topology from the Differentiable Viewpoint*, University Press of Virginia, Charlottesville, 1965.

[34] C. G. T. d. A. Moreira, Hausdorff measures and the Morse-Sard theorem, *Publ. Mat.* **45** (2001), 149–162.

[35] C. G. T. d. A. Moreira and M. A. S. Ruas, The curve selection lemma and the Morse-Sard theorem, *Manuscripta Math.* **45** (2009), 149–162.

[36] A. Morse, The behavior of a function on its critical set, *Ann. of Math.* **40** (1939), 62–70.

[37] M. Morse and A. Sard, *On the Measure of the Critical Values of Functions*, (unpublished), 1935.

[38] A. Norton, A critical set with nonnull image has large Hausdorff dimension, *Trans. Amer. Math. Soc.* **296** (1986), 367–376.

[39] A. Norton, *Ph.D. Thesis*, UC Berkeley Mathematics Department, 1986.

[40] A. Norton, The Zygmund Morse-Sard theorem, *J. Geom. Anal.* **4** (1994), 403–424.

[41] R. van der Putten, The Morse-Sard theorem for Sobolev spaces in a borderline case, *Bull. Sci. Math.* **136** (2012), 463–475.

[42] R. van der Putten, The Morse-Sard theorem in $W^{n,n}(\Omega)$: A simple proof, *Bull. Sci. Math.* **136** (2012), 477–483.

[43] A. Sard, The measure of the critical values of differentiable maps, *Bull. Amer. Math. Soc.* **48** (1942), 883–897.

[44] A. Sard, The equivalence of $n$-measure and Lebesgue in $E_n$, *Bull. Amer. Math. Soc.* **49** (1943), 758–759.

[45] A. Sard, Images of critical sets, *Ann. Math. (2)* **68** (1958), 247–259.

[46] A. Sard, Hausdorff measure of critical images on Banach manifolds, *Amer. J. Math.* **87** (1965), 158–174.

[47] S. Smale, An infinite dimensional version of Sard's theorem, *Amer. J. Math.* **87** (1965), 861–866.

[48] M. Spivak, *Calculus on Manifolds : A Modern Approach to Classical Theorems of Advanced Calculus*, Benjamin, New York, 1965.

[49] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, *Princeton Mathematical Series*, Princeton University Press, Princeton, 1970.

[50] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, *Trans. Am. Math. Soc.* **36** (1934), 63–89.

[51] H. Whitney, A function not constant on a connected set of critical points, *Duke. Math. J.* **1** (1935), 514–517.

[52] Y. Yomdin, The geometry of critical and near-critical values of differentiable mappings, *Math. Ann.* **264** (1983), 495–515.