[version: January 20, 2010]

Structure of level sets and Sard-type properties of Lipschitz maps: results and counterexamples

G. Alberti, S. Bianchini, G. Crippa

Abstract: We consider certain properties of maps of class C^2 from $\mathbb{R}^d$ to $\mathbb{R}^{d-1}$ that are strictly related to Sard's theorem, and show that some of them can be extended to Lipschitz maps, while others still require some additional regularity. We also give counterexamples showing that, in term of regularity, our results are optimal.

Keywords: Lipschitz maps, level sets, singular set, Sard's theorem, weak Sard property, distributional divergence.

Mathematics Subject Classification (2000): 26B35 (26B10, 26B05, 49Q15, 58C25)

1. Introduction

In this paper we prove some results, and construct some counterexamples, concerning three questions that are strictly related to Sard's theorem: structure of the level sets of maps from $\mathbb{R}^d$ to $\mathbb{R}^{d-1}$, weak Sard property of functions from $\mathbb{R}^2$ to $\mathbb{R}$, (strong) locality of the divergence operator.

The original motivation for studying these properties lies in the characterization (given in [1]) of the divergence-free vector fields b on the plane for which the initial value problem for the continuity equation

$$\partial_t u + \operatorname{div}(bu) = 0 \tag{1.1}$$

has a unique bounded solution (in the sense of distributions) for every bounded initial datum.

Structure of level sets. In case of maps $f : \mathbb{R}^d \to \mathbb{R}^{d-1}$ of class C^2, Sard's theorem (see [18] or [13], Theorem 1.3 of Chapter 3)[1] states that the set of singular values of f, namely the image according to f of the singular set

$$S := \{x : \operatorname{rank}(\nabla f(x)) < d-1\}, \tag{1.2}$$

has (Lebesgue) measure zero. By the Implicit Function Theorem, this property implies the following structure result: for a.e. $y \in \mathbb{R}^{d-1}$ the connected components of the level set $E_y := f^{-1}(y)$ are simple curves (of class C^2).

[1] For more general formulations see [6], [9], Theorem 3.4.3, [2], [3].

For Sard's theorem to hold, the regularity assumption on f can be variously weakened, but in any case f must be at least twice differentiable in the Sobolev sense.[2] On the other hand, a variant of the structure theorem holds even with lower regularity: in Theorem 2.5, statement (iv), we prove that if $f : \mathbb{R}^2 \to \mathbb{R}$ is Lipschitz, then for a.e. $y \in \mathbb{R}$ the connected components of E_y are either simple curves or consist of single points (cf. Remark 2.6(ii)), and in statement (v) we prove the same holds if $d \geq 3$ and $f : \mathbb{R}^d \to \mathbb{R}^{d-1}$ is a map of class $C^{1,1/2}$.

Note that for $d \geq 3$ it is not enough to assume that f is Lipschitz: in Section 3 we construct examples of maps $f : \mathbb{R}^d \to \mathbb{R}^{d-1}$ of class $C^{1,\alpha}$ with $\alpha < 1/(d-1)$ and $d \geq 3$ such that for an open set of values y the level set E_y contains a Y-shaped subset, or *triod* (see §2.3 for a precise definition), and therefore at least one connected component of E_y is neither a point nor a curve.[3]

Weak Sard property. Given a function $f : \mathbb{R}^2 \to \mathbb{R}$ of class C^2, consider the measure μ on $\mathbb{R}$ given by the push-forward according to f of the restriction of Lebesgue measure to the singular set S defined in (1.2). The measure μ is supported on the set $f(S)$, which is negligible by Sard's theorem, and therefore μ is singular w.r.t. the Lebesgue measure on $\mathbb{R}$:

$$\mu \perp \mathscr{L}^1 . \tag{1.3}$$

Formula (1.3) can be viewed as a weak version of Sard's theorem, and indeed it holds under the weaker assumption that f is (locally) of class $W^{2,1}$ (cf. [1], Remark ****).

In Section 4 we show that this assumption is essentially optimal; more precisely we construct a function $f : \mathbb{R}^2 \to \mathbb{R}$ of class $C^{1,\alpha}$ for every $\alpha < 1$ (and therefore also of class $W^{\beta,p}$ for every $\beta < 2$ and $p \leq +\infty$) such that (1.3) does not hold.

We actually show that this function does not satisfy a weaker version of (1.3), called *weak Sard property*, where in the definition of μ the singular set S is replaced by $S \cap E^*$, with E^* the union of all connected components with positive length of all level sets. The interest for this property lies in the following result ([1], Theorem ****): given a bounded, divergence-free vector field b on the plane, the equation (1.1) admits a unique bounded solution for every bounded initial datum u_0 if and only if b can be written as $b = (\nabla f)^\perp$ where f is a Lipschitz function with the weak Sard property, and $(\nabla f)^\perp$ is the gradient of f rotated by $90°$ counter-clockwise.

Locality of the divergence operator. It is well-known that given a function u on $\mathbb{R}^d$ which is (locally) of class $W^{1,1}$, the (distributional) gradient ∇u vanishes

[2] It suffices that f is of class $C^{1,1}$, cf. [2], or continuous and (locally) of Sobolev class $W^{2,p}$ with $p > d$, cf. [4], [10]; the constructions in [20], [12] give counterexamples to Sard's Theorem of class $C^{1,\alpha}$ for every $\alpha < 1$, and therefore also of class $W^{\beta,p}$ for every $\beta < 2$ and $p \leq \infty$.

[3] Indeed the key point in the proof of statements (iv) and (v) of Theorem 2.5 is to show that generic level sets contain no triods.

a.e. on every Borel set E where u takes a.e. a constant value. This property is summarized by saying that the gradient is *strongly local* for Sobolev functions; it follows immediately that every first-order differential operator, including the divergence, is strongly local for (first-order) Sobolev functions.

It is then natural to ask whether first-order differential operators are strongly local even on larger spaces.

In Section 5 we show that, somewhat surprisingly, the answer for the divergence operator is negative even in two dimensions,[4] and this fact is strictly related to the (lack of) weak Sard property for Lipschitz functions (see Remark 5.1). More precisely we construct a bounded vector field b on the plane whose (distributional) divergence belongs to L^∞, is non-trivial, and its support is contained in the set where b vanishes. From this vector field we derive another example of Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}$ without the weak Sard property.

To conclude, we point out an interesting open problem concerning Sard's theorem for Sobolev maps: it follows from the results in [4], [10] that Sard's theorem holds for functions $f : \mathbb{R}^2 \to \mathbb{R}$ of class $C^1 \cap W^{2,p}$ with $p > 2$, and the question is whether the same is true for $p \le 2$.

Among other things, a negative answer to this question might provide a negative answer to a question concerning the uniqueness problem for the continuity equation (1.1), and more precisely whether for a continuous, divergence-free vector field b of class $W^{1,p}$,[5] the solution of (1.1) for an arbitrary bounded initial datum u_0 is unique even within the class of *measure-valued* solutions (for more details see [1], ****).

2. Structure of level sets of Lipschitz maps

In this section we state and prove the main result on the structure of level sets of Lipschitz maps (Theorem 2.5). We begin by recalling some basic notation and definitions used through the entire paper, more specific definitions will be introduced when needed.

2.1. Basic notation. - Through the rest of this paper, sets and functions are tacitly assumed to be Borel measurable, and measures are always defined on the appropriate Borel σ-algebra.

Given a subset E of a metric space X, we write $1_E : X \to \{0,1\}$ for the characteristic function of E, $\mathrm{Int}(E)$ for the the interior of E, and, for every $r > 0$, $\mathscr{I}_r E$ for the closed r-neighbourhood of E, that is,

$$\mathscr{I}_r E := \big\{ x \in X : \ \mathrm{dist}(x, E) \le r \big\} .$$

[4] This answers in the negative a question raised by L. Ambrosio; a simpler but less explicit example has been constructed by C. De Lellis and B. Kirchheim.

[5] Note that b satisfies the assumptions of the uniqueness theorem in [5].

The class $\mathscr{F}(X)$ of all non-empty, closed subsets of X is endowed with the *Hausdorff distance*

$$d_H(C, C') := \inf\left\{r \in [0, +\infty] \, : \, C \subset \mathscr{I}_r C', \; C' \subset \mathscr{I}_r C\right\}. \tag{2.1}$$

A function (or a map) defined on a closed set E in $\mathbb{R}^d$ is of class C^k if it admits an extension of class C^k to some open neighbourhood of E, and is of class $C^{k,\alpha}$ with $0 \le \alpha \le 1$ if it is of class C^k and the k-th derivative is Hölder continuous with exponent α.

Given a measure μ on X and a positive function ρ on X we denote by $\rho \cdot \mu$ the measure on X defined by $[\rho \cdot \mu](A) := \int_A \rho \, d\mu$. Hence $1_E \cdot \mu$ is the restriction of μ to the set E.

Given a map $f : X \to X'$ and a measure μ on X, the *push-forward* of μ according to f is the measure $f_\# \mu$ on X' defined by $[f_\# \mu](A) := \mu(f^{-1}(A))$ for every Borel set A contained in X'.

As usual, $\mathscr{L}^d$ is the Lebesgue measure on $\mathbb{R}^d$ while $\mathscr{H}^d$ the d-dimensional Hausdorff measure on every metric space – the usual d-dimensional volume for subsets of d-dimensional surfaces of class C^1 in some Euclidean space. The *length* of a set E is just the 1-dimensional Hausdorff measure $\mathscr{H}^1(E)$. When the measure is not specified, it is assumed to be the Lebesgue measure.

A set E in $\mathbb{R}^d$ is *k-rectifiable* if it can be covered, except for an $\mathscr{H}^k$-negligible subset, by countably many k-dimensional surfaces of class C^1.

2.2. CURVES. - A *curve* in $\mathbb{R}^d$ is the image C of a continuous, non-constant path $\gamma : [a, b] \to \mathbb{R}^d$ (the parametrization of C); thus C is a compact connected set that contains infinitely many points. We say that C is *simple* if it admits a parametrization γ that is injective, *closed* if γ satisfies $\gamma(a) = \gamma(b)$; and *closed and simple* if $\gamma(a) = \gamma(b)$ and γ is injective on $[a, b)$.[6]

If γ is a parametrization of C of class $W^{1,1}$, then $\mathscr{H}^1(C) \le \|\dot\gamma\|_1$ and the equality holds whenever γ is injective. Moreover it is always possible to find a strictly increasing function $\sigma : [a', b'] \to [a, b]$ such that $\gamma \circ \sigma$ is a Lipschitz parametrization which satisfies $|(\gamma \circ \sigma)'| = 1$ a.e..

For closed curves, it is sometimes convenient to identify the end points of the domain $[a, b]$. This quotient space is denoted by $[a, b]^*$, and endowed with the distance

$$d(x, y) := \min\{|x - y|, (b - a) - |x - y|\}. \tag{2.2}$$

A set E in $\mathbb{R}^d$ is *path-connected* if every couple of points $x, y \in E$ can be joined by a curve contained in E (that is, x, y agree with the end points $\gamma(a), \gamma(b)$ of the curve).

[6] The usual definitions of curve and length may differ from ours, even though they are ultimately equivalent. For more details, see [8], Section 3.2. For a detailed study of spaces with finite length, see [11].

2.3. Triods. - A *(simple) triod* in $\mathbb{R}^d$ is any set Y given by the union of three curves with only one end point y in common, which we call *center* of Y. More precisely

$$Y = C_1 \cup C_2 \cup C_3$$

where each C_i is a curve in $\mathbb{R}^d$ parametrized by $\gamma_i : [a_i, b_i] \to C_i$ and $\gamma_i(a_i) = y$, and the sets $\gamma_i((a_i, b_i])$ are pairwise disjoint.

2.4. Lipschitz maps. - Through this section d, k are positive integers such that $0 < k < d$, and f is a Lipschitz map from (a subset of) $\mathbb{R}^d$ to $\mathbb{R}^{d-k}$; we denote by f_i, $i = 1, \dots, d-k$, the components of f.

For every $y \in \mathbb{R}^{d-k}$ we denote by $E_y := f^{-1}(y)$ the corresponding level set of f, and by $\mathscr{C}_y$ the family of all connected components C of E_y such that $\mathscr{H}^k(C) > 0$;[7] E^*_y is the union of all C in $\mathscr{C}_y$, and E^* is the union of all E^*_y with $y \in \mathbb{R}^{d-k}$. Both E^*_y and E^* are Borel sets (Proposition 6.1).

For every point x where f is differentiable it is possible to define the Jacobian

$$Jf := [\det(\nabla f \cdot \nabla^t f)]^{1/2} .$$

If $Jf(x) \neq 0$ the matrix $\nabla f(x)$ has rank $d-k$. Thus, if $k = 1$ and $Jf(x) \neq 0$, there exists a unique unit vector $\tau = \tau(x)$ such that τ is orthogonal to ∇f_i for every i and the sequence $(\nabla f_1, \dots, \nabla f_{d-1}, \tau)$ is a positively oriented basis of $\mathbb{R}^d$.[8]

2.5. Theorem. - *Let $f : \mathbb{R}^d \to \mathbb{R}^{d-k}$ be a Lipschitz map with compact support. In the notation of the previous paragraphs, the following statements hold for almost every $y \in \mathbb{R}^{d-k}$:*

(i) *the level set E_y is k-rectifiable and $\mathscr{H}^k(E_y) < +\infty$;*

(ii) *for $\mathscr{H}^k$-a.e. $x \in E_y$ the map f is differentiable at x, the matrix $\nabla f(x)$ has rank $d-k$, and its kernel is the tangent space to E_y at x;*

(iii) *the family $\mathscr{C}_y$ is countable and $\mathscr{H}^k(E_y \setminus E^*_y) = 0$;*

(iv) *for $k = 1$ and $d = 2$ every connected component C of E_y is either a point or a closed simple curve with a Lipschitz parametrization $\gamma : [a, b]^* \to C$ which is injective and satisfies $\dot\gamma(t) = \tau(\gamma(t))$ for a.e. t;*

(v) *the result in the previous statement can be generalized to $k = 1$ and $d \geq 3$ provided that f is of class $C^{1,1/2}$.*

2.6. Remarks. - (i) The assumption that f is defined on $\mathbb{R}^d$ and has compact support was made for the sake of simplicity. Under more general assumptions, the results on the local properties of generic level sets (rectifiability and so on) are clearly the same, while the results concerning the global structure (notably statement (iv)) require some obvious modifications.

[7] Since the length of a connected set is larger than its diameter, for $k = 1$ the connected components in $\mathscr{C}_y$ are just those that contain more than one point.

[8] If $d = 2$ and $k = 1$ then $Jf = |\nabla f|$ and τ is the counter-clockwise rotation by 90° of $\nabla f / |\nabla f|$.

(ii) In statement (iv), the possibility that some connected components of E_y are points cannot be ruled out. Indeed, for every $\alpha < 1$, there are functions $f : \mathbb{R}^2 \to \mathbb{R}$ of class $C^{1,\alpha}$ such that the set $E_y \setminus E_y^*$, namely the union of all connected components of E_y which consist of single points (cf. footnote 7), is not empty for an interval of values y.[9]

(iii) In Section 3 we show that for every $d \geq 3$ and every $\alpha < 1/(d-1)$ there exist maps $f : \mathbb{R}^d \to \mathbb{R}^{d-1}$ of class $C^{1,\alpha}$ such that the level set E_y contains a simple triod (cf. §2.3) for an open set of values y, and therefore its connected components cannot be just points or simple curves. This shows that the exponent $1/2$ in the assumption $f \in C^{1,1/2}$ of statement (v) is optimal for $d = 3$. However, this exponent seems to be specific of our proof, and we believe that it is not optimal for $d \geq 4$.

(iv) The fact that a connected component of E_y is a simple curve with Lipschitz parametrization (statement (iv)) does not imply that it can be locally represented as the graph of a Lipschitz function: there exist functions f on the plane even of class $C^{1,\alpha}$ with $\alpha < 1$ such that every level set contains a cusp.

(v) Even knowing that a connected component of E_y is a simple curve, the existence of a parametrization whose velocity field agrees a.e. with τ (cf. statement (iv)) is not as immediate as it may look, because τ and E_y lack almost any regularity. Our proof relies on the fact that generic level sets of Lipschitz maps can be endowed with the structure of rectifiable currents without boundary.

(vi) One might wonder if something similar to statement (iv) holds at least for functions $f : \mathbb{R}^d \to \mathbb{R}$ with $d > 2$. As shown below, the key points in the proof of statement (iv) are that a family of pairwise disjoint triods in the plane is countable, and that a connected set in the plane with finite length which contains no triods is a simple curve. A natural generalization of the notion of triod could be the following: a connected, compact set E in $\mathbb{R}^d$ is a d-triod with center y if, for every open ball B which contains y, the set $\overline{B} \setminus E$ has at least three connected components which intersect ∂B. However, even if it is still true that a family of pairwise disjoint d-triods in $\mathbb{R}^d$ is countable, very little can be said on the topological structure of d-triod-free connected sets with finite $\mathscr{H}^{d-1}$ measure.

The rest of this section is devoted to the proof of Theorem 2.5.

2.7. Coarea formula. - The coarea formula (see e.g. [9], §3.2.11, or [19], §10, or [15], Corollary 5.2.6) states that for every Lipschitz map $f : \mathbb{R}^d \to \mathbb{R}^{d-k}$

[9] Simple examples can be obtained by modifying the construction in [12]. Moreover, since the level sets E_y of any continuous function contain isolated points only for countably many y, it turns out that $E_y \setminus E_y^*$ has the cardinality of continuum for a set of positive measure of y. An example of Lipschitz function on the plane such that the set $E_y \setminus E_y^*$ has the cardinality of continuum for a.e. y is also given in [14].

and every positive Borel function $\phi : \mathbb{R}^d \to [0, +\infty]$ there holds

$$\int_{\mathbb{R}^d} \phi \, Jf \, d\mathscr{L}^d = \int_{\mathbb{R}^{d-k}} \left[\int_{E_y} \phi \, d\mathscr{H}^k \right] d\mathscr{L}^{d-k}(y) \tag{2.3}$$

(quite obviously, (2.3) holds even for real-valued functions ϕ provided that the integral at the left-hand side makes sense, e.g., for $\phi \in L^1(\mathbb{R}^d)$). The coarea formula has two immediate consequences:

(i) Let ϕ be the constant function 1: if f has compact support then the integral at the left-hand side of (2.3) is finite, and therefore $\mathscr{H}^k(E_y)$, which is the value of the integral between square brackets at the right-hand side, is finite for a.e. $y \in \mathbb{R}^{d-k}$.

(ii) Let S be the set of all points x in $\mathbb{R}^d$ where either f is not differentiable or $Jf(x) = 0$. If ϕ is the characteristic function of S, then identity (2.3) implies that for a.e. $y \in \mathbb{R}^{d-k}$ there holds $\mathscr{H}^k(E_y \cap S) = 0$. This means that for $\mathscr{H}^k$-a.e. $x \in E_y$ the map f is differentiable at x and $Jf(x) \neq 0$, that is, $\nabla f(x)$ has rank $d - k$. Hence E_y admits a k-dimensional tangent space at x, namely the kernel of the matrix $\nabla f(x)$. This implies that E_y is k-rectifiable (cf. [19], Theorem 10.4, or [9], Theorem 3.2.15).

Proof of statements (i) and (ii) of Theorem 2.5. - These statements are contained in §2.7. □

Statement (iii) of Theorem 2.5 will be obtained as a corollary of Lemmas 2.11, 2.12, and 2.13. In the next three paragraphs we recall some definitions and a few results that will be used in the proofs of these lemmas.

2.8. Connected components. - We recall here some basic facts about the connected components of a set E in $\mathbb{R}^d$, or more generally in a metric space X; for more details see for instance [7], Chapter 6.

A *connected component* of E is any element of the class of connected subsets of E which is maximal with respect to inclusion. The connected components of E are pairwise disjoint, closed in E, and cover E.

Assume now that E is compact. Then each connected component C agrees with the intersection of all subsets of E that are closed and open in E and contain C ([7], Theorem 6.1.23).

Every set which is open and closed in E can be written as $U \cap E$ where U is an open subset of X such that $\partial U \cap E = \varnothing$.[10] Hence C is the intersection of the closures of all open sets U which contain C and satisfy $\partial U \cap E = \varnothing$. Therefore, under the further assumption that X is second countable (or, equivalently,

[10] If D is open and closed in E, then D and $E \setminus D$ are disjoint and closed in X, and since X is a normal space there exist disjoint open sets U, V such that $D \subset U$ and $E \setminus D \subset V$; it is then easy to checks that U meets all requirements.

separable), we can find a decreasing sequence U_n of such sets, the intersection of whose closures is still C.[11]

2.9. RECTIFIABLE CURRENTS. - We recall here some basic definitions about currents; for further details see for instance [19], Chapter 6, or [15], Chapter 7.

A k-dimensional *current* on $\mathbb{R}^d$ is a linear functional on the space of k-forms on $\mathbb{R}^d$ of class C_c^∞. The *boundary* of a k-current T is the $(k-1)$-dimensional current ∂T defined by $\langle \partial T; \omega \rangle := \langle T; d\omega \rangle$, where $d\omega$ is the exterior derivative of ω. The mass of T, denoted by $\mathbb{M}(T)$, is the supremum of $\langle T; \omega \rangle$ among all k-forms ω that satisfy $|\omega| \leq 1$ everywhere; $\mathbb{M}$ is clearly a norm on the subspace of currents with finite mass.

Let E be a k-rectifiable set in $\mathbb{R}^d$. An *orientation* of E is a map ζ that associates to $\mathscr{H}^k$-a.e. x in E a unit, simple k-vector that spans the approximate tangent space to E at x; a multiplicity is any integer-valued function m on E which is locally summable w.r.t. $\mathscr{H}^k$. To every choice of E, ζ, m is canonically associated the k-dimensional current $[E, \zeta, m]$ defined by

$$\langle [E, \zeta, m]; \omega \rangle := \int_E \langle \omega; \zeta \rangle \, m \, d\mathscr{H}^k$$

for every k-form ω of class C_c^∞ on $\mathbb{R}^d$. Hence the mass of $[E, \zeta, m]$ is equal to $\int_E |m| \, d\mathscr{H}^k$. Currents of this type are called *rectifiable*.[12] .

2.10. CURRENT STRUCTURE OF LEVEL SETS. - For every $x \in \mathbb{R}^d$ where f is differentiable and the matrix ∇f has rank $d-k$, we choose an orthonormal basis $\{\tau_1, \dots, \tau_k\}$ of the kernel of ∇f so that the sequence $(\nabla f_1, \dots, \nabla f_{d-k}, \tau_1, \dots, \tau_k)$ is a positively oriented basis of $\mathbb{R}^d$. By §2.7(ii) the k-vector field $\tau := \tau_1 \wedge \ldots \wedge \tau_k$ is an orientation of E_y for a.e. y.[13]

We then denote by T_y the k-dimensional current associated with the set E_y, the orientation τ, and constant multiplicity 1, that is, $T_y := [E_y, \tau, 1]$. The essential fact about the current T_y is that its boundary vanishes for a.e. y, that is

$$\langle \partial T_y; \omega \rangle := \langle T_y; d\omega \rangle := \int_{E_y} \langle d\omega; \tau \rangle \, d\mathscr{H}^k = 0 \tag{2.4}$$

for every $(k-1)$-form ω of class C_c^∞ on $\mathbb{R}^d$. The proof goes as follows: the currents T_y are the slices according to the map f of the d-dimensional rectifiable

[11] Recall that in a second countable space every family $\mathscr{F}$ of closed sets admits a countable subfamily $\mathscr{F}'$ such that the intersection of $\mathscr{F}'$ and the intersection of $\mathscr{F}$ agree.

[12] Examples of k-dimensional rectifiable currents are obtained by taking a k-dimensional oriented surface E of class C^1, endowed with constant multiplicity 1. In this case the mass agrees with the k-dimensional volume of E, and the boundary of E in the sense of currents agrees with the $(k-1)$-dimensional current associated to the usual boundary ∂E, endowed with the canonical orientation and constant multiplicity 1.

[13] For $k = 1$, τ agrees with the vector field defined in §2.4.

current $T := [\mathbb{R}^d, \zeta, 1]$, where ζ is the canonical orientation of $\mathbb{R}^d$. In general, the boundaries of the slices T_y agree with the slices of the boundary ∂T for a.e. y (see [9], §4.3.1), and since in this particular case $\partial T = 0$, then $\partial T_y = 0$ for a.e. y.

2.11. Lemma. - *Let $T := [E, \zeta, m]$ be a rectifiable k-current in $\mathbb{R}^d$, and let $T' := [E \cap A, \zeta, m]$ where A is a set in $\mathbb{R}^d$. If $\partial T = 0$ and the boundary of A does not intersect the closure of E, then $\partial T' = 0$.*

Proof. - Since $\partial A \cap \overline{E} = \varnothing$, the sets $E \cap A$ and $E \setminus A$ have disjoint closures, and we can find a smooth function $\sigma : \mathbb{R}^d \to [0, 1]$ that is equal to 1 on some neighbourhood of $E \cap A$, and to 0 on some neighbourhood of $E \setminus A$. Then, for every form ω there holds[14]

$$\begin{aligned} \langle \partial T'; \omega \rangle = \langle T'; d\omega \rangle = \langle T; \sigma\, d\omega \rangle &= \langle T; \sigma\, d\omega + d\sigma \wedge \omega \rangle \\ &= \langle T; d(\sigma\omega) \rangle = \langle \partial T; \sigma\omega \rangle = 0 \,. \end{aligned}$$ □

2.12. Lemma. - *Let $T := [E, \zeta, m]$ be a rectifiable k-current in $\mathbb{R}^d$ with E bounded, and let $T' := [E \cap C, \zeta, m]$ where C is a connected component of the closure of E. If $\partial T = 0$ then $\partial T' = 0$.*

Proof. - Since C is a connected component of $\overline{E}$, we can find a decreasing sequence of open sets U_n such that the intersection is C and $\partial U_n \cap \overline{E} = \varnothing$ for every n (see §2.8). We set $T_n := [E \cap U_n, \zeta, m]$. Hence Lemma 2.11 implies $\partial T_n = 0$, and since T_n converge to T',[15] then $\partial T' = 0$. □

2.13. Lemma. - *Let E_y be a level set of f such that $\mathscr{H}^k(E_y) < +\infty$ and the associated current T_y is well-defined and has no boundary (cf. §2.10). Then $\mathscr{H}^k(E_y \setminus E_y^*) = 0$.*

Proof. - Set $B := E_y \setminus E_y^*$, fix $\delta > 0$, and take an open set A_δ such that $B \subset A_\delta$ and

$$\mathscr{H}^k(E_y \cap A_\delta) \leq \mathscr{H}^k(B) + \delta \,.$$

Recall that B is the union of the connected components of E_y which are $\mathscr{H}^k$-negligible. For every such connected component C, we can find an open neighbourhood U such that $C \subset U \subset A_\delta$, $\mathscr{H}^k(E_y \cap U) \leq \delta$, and $\partial U \cap E_y = \varnothing$ (see §2.8). From such family of neighbourhoods we extract a countable subfamily $\{U_n\}$ that covers B, and set $V_n := U_n \setminus (U_1 \cup \ldots \cup U_{n-1})$ for every n. The sets V_n are pairwise disjoint and cover B, and one easily checks that $\partial V_n \cap E_y = \varnothing$ for every n.

We then set $T_n := [E_y \cap V_n, \tau, 1]$. Thus the sum $\sum_n T_n$ agrees with the current $T_\delta := [E_y \cap V_\delta, \tau, 1]$ where V_δ is the union of the sets V_n. Moreover the properties

[14] The second identity follows by the definition of T' and the fact that $\sigma = 1$ on $E \cap A$ and $\sigma = 0$ on $E \setminus A$. The third identity follows by the fact that $d\sigma = 0$ on E.

[15] We have $\mathbb{M}(T_n - T) \to 0$, which implies convergence in the sense of currents.

of U_n and V_n yield $\mathbb{M}(T_n) = \mathscr{H}^k(E_y \cap V_n) \le \mathscr{H}^k(E_y \cap U_n) \le \delta$ and $\partial T_n = 0$ (apply Lemma 2.11). Thus the isoperimetric theorem (see [19], Theorem 30.1, or [15], Theorem 7.9.1) yields $T_n = \partial S_n$ for some rectifiable $(k+1)$-current S_n that satisfies

$$\mathbb{M}(S_n) \le c\,[\mathbb{M}(T_n)]^{1+1/k} \le c\,\mathbb{M}(T_n)\,\delta^{1/k}\,,$$

where c is a constant that depends only on d and k. Since the *flat norm*[16] of T_n satisfies $\mathbb{F}(T_n) \le \mathbb{M}(S_n)$, the previous inequality yields

$$\mathbb{F}(T_n) \le c\,\mathbb{M}(T_n)\,\delta^{1/k} = c\,\mathscr{H}^k(E_y \cap V_n)\,\delta^{1/k}\,.$$

Taking the sum over all n we obtain $\mathbb{F}(T_\delta) \le c\,\mathscr{H}^k(E_y)\,\delta^{1/k}$, and therefore T_δ tends to 0 w.r.t. the flat norm as $\delta \to 0$.

On the other hand, the inclusions $B \subset V_\delta \subset A_\delta$ and the choice of A_δ imply $\mathscr{H}^k((E_y \cap V_\delta) \setminus B) \le \delta$. Hence $T_\delta := [E_y \cap V_\delta, \tau, 1]$ converge to the current $[B, \tau, 1]$ w.r.t. the norm $\mathbb{M}$, and therefore also w.r.t. the flat norm $\mathbb{F}$ as $\delta \to 0$. Thus $[B, \tau, 1]$ must be equal to 0, and this means $\mathscr{H}^k(B) = 0$. □

Proof of statement (iii) of Theorem 2.5. - Let A be the set of all y such that statements (i) and (ii) of Theorem 2.5 hold and $T_y := [E_y, \tau, 1]$ is a well-defined current without boundary (cf. §2.10). Then A has full measure in $\mathbb{R}^{d-k}$, and we claim that statement (iii) holds for every $y \in A$. Indeed, the elements of $\mathscr{C}_y$ are pairwise disjoint subsets of E_y with positive $\mathscr{H}^k$-measure, and since $\mathscr{H}^k(E_y)$ is finite, $\mathscr{C}_y$ must be countable. Moreover $\mathscr{H}^k(E_y \setminus E_y^*) = 0$ by Lemma 2.13. □

Next we give some lemmas used in the proof of statements (iv) and (v) of Theorem 2.5.

2.14. Lemma. - *Let $T := [C, \zeta, 1]$ be a 1-dimensional rectifiable current in $\mathbb{R}^d$, where C is a curve with Lipschitz parametrization $\gamma : [a,b] \to C$ s.t. $|\dot\gamma| = 1$ a.e. Assume that $\partial T = 0$ and C is simple. Then*

(i) *C is closed;*

(ii) *either $\zeta \circ \gamma = \dot\gamma$ a.e. in $[a,b]$ or $\zeta \circ \gamma = -\dot\gamma$ a.e. in $[a,b]$.*

Proof. - Since $\zeta(\gamma(t))$ and $\dot\gamma(t)$ are parallel unit vectors for a.e. t, there exists $\sigma : [a,b] \to \{\pm 1\}$ such that $\zeta(\gamma(t)) = \sigma(t)\,\dot\gamma(t)$ for a.e. t. Thus statement (ii) amounts to say that σ is a.e. constant.

Step 1. Since γ is injective at least on $[a,b)$, the assumption $\partial T = 0$ can be re-written as

$$0 = \int_C \langle d\phi; \zeta\rangle\, d\mathscr{H}^1 = \int_a^b \langle d\phi \circ \gamma; \dot\gamma\rangle\, \sigma\, d\mathscr{L}^1$$

[16] Here the flat norm $\mathbb{F}(T)$ of a current T with compact support is the infimum of $\mathbb{M}(R) + \mathbb{M}(S)$ over all possible currents R, S such that $T = R + \partial S$, cf. [9], §4.1.12.

for every function (0-form) ϕ on $\mathbb{R}^d$ of class C_c^∞. Since $\langle d\phi \circ \gamma; \dot\gamma \rangle$ is the (distributional) derivative of $\phi \circ \gamma$, we obtain that

$$0 = \int_a^b \dot\varphi \, \sigma \, d\mathscr{L}^1 \tag{2.5}$$

for every test function $\varphi : [a,b] \to \mathbb{R}$ of the form $\varphi = \phi \circ \gamma$ where ϕ is a function on $\mathbb{R}^d$ of class C_c^∞.

Step 2. We assume first that $\gamma(a) = \gamma(b)$ and γ is injective on the interval with identified end points $[a,b]^*$, that is, C is simple and closed (cf. §2.2).

By density we can immediately extend identity (2.5) to all test function φ of the form $\varphi = \phi \circ \gamma$ with ϕ Lipschitz, and these are exactly the functions ϕ that are Lipschitz w.r.t. the distance d' defined by $d'(x,y) := |\gamma(x) - \gamma(y)|$ for every $x, y \in [a,b]^*$.

Now, it follows from Propositions 7.8 and 7.10 (cf. Remark 7.11) that every function φ on $[a,b]^*$ which is Lipschitz w.r.t. the distance d defined in (2.2) can be uniformly approximated by functions that are uniformly Lipschitz w.r.t. d and Lipschitz w.r.t. d'. This is enough to conclude that (2.5) holds for all $\varphi : [a,b]^* \to \mathbb{R}$ that are Lipschitz w.r.t. d, that is, all $\varphi : [a,b] \to \mathbb{R}$ that are Lipschitz w.r.t. the usual Euclidean distance and satisfies $\varphi(a) = \varphi(b)$. This implies that the distributional derivative of σ vanishes, and therefore σ is a.e. constant.

Step 3. Assume by contradiction that γ is injective on $[a,b]$, that is, C is simple but not closed. Proceeding as in Step 2 we obtain that (2.5) holds for all Lipschitz test function $\varphi : [a,b] \to \mathbb{R}$, without any restrictions on the values at a and b. This allows us to conclude that σ is a.e. null, in contradiction with the fact that it takes values ± 1. □

2.15. Lemma. - *Let $\mathscr{F}$ be a family of pairwise disjoint triods in $\mathbb{R}^2$. Then $\mathscr{F}$ is countable.*

The above lemma has been proved in [16], Theorem 1, (see also [17], Theorem 1). For the reader's convenience, we give a self-contained proof of this result in Section 8.

2.16. Lemma. - *Given a map $f : \mathbb{R}^d \to \mathbb{R}^{d-1}$, the following statements hold:*

(i) *if $d = 2$, then the level set E_y contains no triods for all $y \in \mathbb{R}$ except countably many;*[17]

(ii) *if $d \geq 3$ and f is of class $C^{1,1/2}$, then the level set E_y contains no triods for a.e. $y \in \mathbb{R}^{d-1}$.*

Proof. - Since the level sets of any map are pairwise disjoint, statement (i) follows immediately from Lemma 2.15.

[17] Note that f does not need to be Lipschitz, and not even continuous.

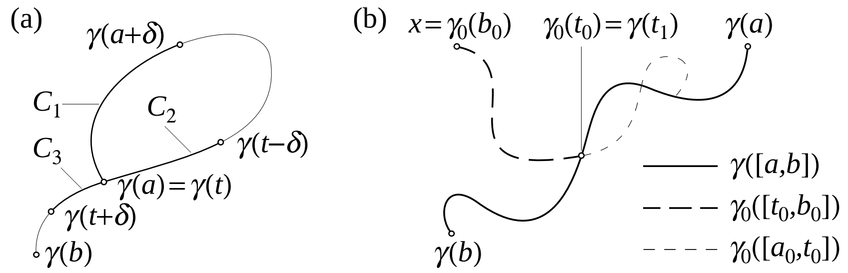

Figure 1. Construction of triods in the proof of Lemma 2.17: (a) Step 3, and (b) in Step 4.

We prove statement (ii) by reduction to the case $d = 2$. Consider the open set

$$U := \left\{ x \in \mathbb{R}^d : \operatorname{rank}(\nabla f(x)) \geq d-2 \right\}.$$

Since f is of class $C^{1,1/2}$, a refined version of Sard theorem ([2], Theorem 2) shows that the level set E_y is contained in U for a.e. $y \in \mathbb{R}^{d-1}$, and therefore it is sufficient to prove statement (ii) for the restriction of f to U.

Now, by the Implicit Function Theorem we can cover U by open sets V where $d-2$ of the variables x_i can be written in terms of $d-2$ of the variables y_i, and more precisely there exist an open set W in $\mathbb{R}^d$, a diffeomorphism $\psi : W \to V$ of class C^1, and a map $g : W \to \mathbb{R}$ such that, after a suitable re-numbering of the variables,

$$f \circ \psi(t) = (t_1, \dots, t_{d-2}, g(t)) \quad \text{for all } t \in W.$$

Therefore it suffices to prove statement (ii) for the map $\tilde{f} := f \circ \psi$.

Let N be the set of all $y \in \mathbb{R}^{d-1}$ such that the level set $\tilde{E}_y$ of $\tilde{f}$ contains a triod, and for every $y' = (y_1, \dots, y_{d-2}) \in \mathbb{R}^{d-2}$ let $N_{y'}$ be the set of all $y'' \in \mathbb{R}$ such that $(y', y'') \in N$. Then statement (i) shows that $N_{y'}$ is countable for every y', and therefore Fubini's Theorem implies that N is negligible. □

2.17. Lemma. - *Let E be a closed, connected set in $\mathbb{R}^d$ with finite, strictly positive length. If E contains no triods, then it is a simple curve, possibly closed. More precisely, there exists a Lipschitz parametrization $\gamma : [a,b] \to E$ which is either injective on $[a,b]$ or satisfies $\gamma(a) = \gamma(b)$ and is injective on $[a,b)$.*

Proof. - *Step 1.* Recall the following well-known fact: a connected closed set E with finite length is connected by simple curves. More precisely, for every x, y in E there exists an *injective* Lipschitz map $\gamma : [a,b] \to C$ such that $\gamma(a) = x$ and $\gamma(b) = y$ (see for instance [8], Lemma 3.12).

Step 2. Let $\mathscr{F}$ be the family of all Lipschitz maps $\gamma : [a,b] \to E$ (the domain $[a,b]$ may vary with γ) that are injective on (a,b) and satisfy $|\dot{\gamma}| \leq 1$ a.e. We order $\mathscr{F}$ by inclusion of graphs, that is, $\gamma_1 \preceq \gamma_2$ if $[a_1,b_1] \subset [a_2,b_2]$ and $\gamma_1 = \gamma_2$ on $[a_1,b_1]$. One easily checks that $\mathscr{F}$ admits a maximal element $\gamma : [a,b] \to E$. In the next steps we show that this is the parametrization we are looking for.

Step 3. Either γ is injective on $[a, b]$ or satisfies $\gamma(a) = \gamma(b)$ and is injective on $[a, b)$. Since γ is injective on (a, b), it suffices to show that for every $t \in (a, b)$ there holds $\gamma(t) \neq \gamma(a), \gamma(b)$. Assume by contradiction that $\gamma(t) = \gamma(a)$ for some t, and take a positive δ such that $2\delta < t - a$ and $\delta < b - t$. Then, contrary to the assumptions of the statement, E contains the triod with center $y := \gamma(a) = \gamma(t)$ given by the union of the following three curves: $C_1 := \gamma([a, a + \delta])$, $C_2 := \gamma([t - \delta, t])$, $C_3 := \gamma([t, t + \delta])$ – see Figure 1(a).

Step 4. The image of γ is E. Assume by contradiction that there exists $x \in E \setminus \gamma([a, b])$. By Step 1 there exists an injective Lipschitz map $\gamma_0 : [a_0, b_0] \to E$ such that $\gamma_0(b_0) = x$ and $\gamma_0(a_0)$ is some point in $\gamma([a, b])$. We can also assume that $|\dot\gamma_0| \leq 1$ a.e. Now, let t_0 be the largest of all $t \in [a_0, b_0]$ such that $\gamma_0(t) \in \gamma([a, b])$ and take $t_1 \in [a, b]$ so that $\gamma(t_1) = \gamma_0(t_0)$.

If γ is injective on $[a, b]$ – the other case is similar – we derive a contradiction in each of the following cases: i) $t_1 = b$, ii) $t_1 = a$, iii) $a < t_1 < b$. If $t_1 = b$, we extend the map γ by setting $\gamma(t) := \gamma_0(t - b + t_0)$ for all $t \in [b, b + b_0 - t_0]$; one easily checks that the extended map belongs to $\mathscr{F}$, in contradiction with the maximality of γ. A similar contradiction is obtained if $t_1 = a$. Finally, if $a < t_1 < b$ then E contains the triod with center $y := \gamma(t_1) = \gamma_0(t_0)$ given by the union of the following three curves: $C_1 := \gamma_0([t_0, b_0])$, $C_2 := \gamma([a, t_1])$, $C_3 := \gamma([t_1, b])$ – see Figure 1(b). □

Proof of statements (iv) and (v) of Theorem 2.5. - Let $d = 2$ or $d \geq 3$ and f be of class $C^{1,1/2}$. Then for a.e. $y \in \mathbb{R}^{d-1}$ the level set E_y has finite length (§2.7), contains no triods (Lemma 2.16), and the associated current $T_y := [E_y, \tau, 1]$ is well-defined, rectifiable, and without boundary (§2.10).

For every such y, let C be a connected component of E_y which is not a point. Since C contains no triods, Lemma 2.17 implies that C is a simple curve with a Lipschitz parametrization γ, and we can further assume that $|\dot\gamma| = 1$ a.e.

By Lemma 2.12 the current $T' := [C, \tau, 1]$ has no boundary, and therefore Lemma 2.14 implies that C is a closed curve such that either $\tau \circ \gamma = \dot\gamma$ a.e. or $\tau \circ \gamma = -\dot\gamma$ a.e.; in the latter case we conclude the proof by replacing the parametrization $\gamma(t)$ with $\gamma(-t)$. □

3. Examples of maps without triod-free level sets

In this section we show that the assumption that f is of class $C^{1,1/2}$ in statement (v) of Theorem 2.5 cannot be dropped. More precisely, given $d \geq 3$ and $\alpha < 1/(d-1)$, we construct a map f of class $C^{1,\alpha}$ from $\mathbb{R}^d$ to a cube Q_0 in $\mathbb{R}^{d-1}$ such that *every* level set contains a triod,[18] and therefore at least one of its

[18] At first glance, this claim seems to contradict the fact that a generic level set of a Lipschitz map is a rectifiable current with multiplicity 1 and no boundary (cf. §2.10). It is not so, since we do not claim that the level set *coincides* with the triod in a neighbourhood of the center of the triod.

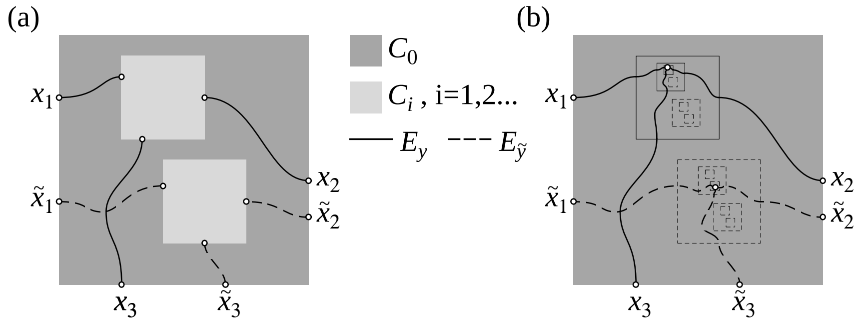

Figure 2. The components of the level sets E_y and $E_{\tilde{y}}$ starting from the points x_j and $\tilde{x}_j$ $(j = 1, 2, 3)$ after the first step of the construction (a) and in the end (b).

connected components is neither a point nor a simple curve (Proposition 3.7(iii)).

This example shows that the Hölder exponent $1/2$ in Theorem 2.6(v) is optimal for $d = 3$. As pointed out in Remark 2.6(iii), we believe the the optimal Hölder exponent for $d > 3$ is the one suggested by this example, namely $1/(d-1)$, and not $1/2$.

3.1. Idea of the construction. - Assume for simplicity that $d = 3$. The strategy for the construction of f is roughly the following: we divide the target square Q_0 in a certain number N of sub-squares Q_i with side-length ρ, then we define f on the cube C_0 minus N disjoint sub-cubes C_i with side-length r, so that the following key property holds: for every $y \in Q_0$ the level set $f^{-1}(y)$ contains three disjoint curves connecting three points x_1, x_2, x_3 on the boundary of C_0 with three points on the boundary of the cube C_i, where the index i is such that y belongs to Q_i, see Figure 2(a).[19]

In the second step we replicate this construction within each C_i, so that f takes values in Q_i, is defined outside N sub-cubes with side-length r^2, and now $f^{-1}(y)$ contains disjoint curves connecting x_1, x_2, x_3 to three points on the boundary of one of these smaller cubes. And so on ...

As one can see from Figure 2(b), in the end each level set E_y will contain three disjoint curves connecting x_1, x_2, x_3 to the same point, which means that E_y contains a triod (cf. §2.3).

Note that in the $(n + 1)$-th step we define f on each one of the N^n cubes with side-length r^n left over from the previous step minus N sub-cubes with side-length r^{n-1}, so that each cube is mapped in a sub-square of Q_0 with side-length ρ^n. Thus the oscillation of f on this cube is ρ^n, and this is enough to guarantee that in the end the map f is continuous. As we shall see, if things are carefully arranged, f turns out to be of class $C^{1,\alpha}$.

[19] This is the only part of this construction that requires more than two dimensions: as one can easily see in Figure 2, in two dimensions the joining curves for different values of y would necessarily intersect, in contradiction with the fact that they are contained in different level sets.

3.2. NOTATION. - We denote the points in $\mathbb{R}^d$ by $x = (x', x_d)$ with $x' := (x_1, \dots, x_{d-1})$, and the points in $\mathbb{R}^{d-1}$ by the letter y. For every $x_0 \in \mathbb{R}^d$ and every $\ell > 0$, we denote by $C(x_0, \ell)$ the closed cube in $\mathbb{R}^d$ with center x_0 and side-length ℓ given by

$$C(x_0, \ell) := x_0 + \left[-\frac{\ell}{2}, \frac{\ell}{2}\right]^d .$$

We denote by $Q(y_0, \ell)$ the closed cube in $\mathbb{R}^{d-1}$, similarly defined.

Through this section we reserve the letter C for d-dimensional closed cubes in $\mathbb{R}^d$ of the form $C(x_0, \ell)$, and the letter Q for the $(d-1)$-dimensional closed cubes in $\mathbb{R}^{d-1}$ or in $\mathbb{R}^d$ with axes parallel to the coordinate axes. With the case $d = 3$ in mind, we often refer to the former ones simply as "cubes" and to the latter ones as "squares".

We set

$$C_0 := C(0,1) = \left[-\frac{1}{2}, \frac{1}{2}\right]^d , \quad Q_0 := Q(0,1) = \left[-\frac{1}{2}, \frac{1}{2}\right]^{d-1} ,$$

and for every cube $C = C(x_0, \ell)$ we denote by g_C is the homothety on $\mathbb{R}^d$ which maps C_0 into C, that is,

$$g_C(x) := x_0 + \ell x \quad \text{for all } x \in \mathbb{R}^d.$$

Similarly, for every square $Q = Q(y_0, \ell)$ we set

$$h_Q(y) := y_0 + \ell y \quad \text{for all } y \in \mathbb{R}^{d-1}.$$

3.3. A COMPLETE NORM FOR $C^{k,\alpha}$. - Given a real or vector-valued map f defined on a subset E of $\mathbb{R}^d$ and $\alpha \in (0,1]$, the homogeneous Hölder (semi)norm of exponent α of f is

$$\|f\|_{C^{0,\alpha}_{\mathrm{om}}} := \sup_{\substack{x,y \in E \\ x \neq y}} \frac{|f(x) - f(y)|}{|x-y|^\alpha} . \tag{3.1}$$

Let be given an open set A in $\mathbb{R}^d$, a point $x_0 \in A$, and an integer $k \geq 0$. Among the many (equivalent) complete norms on the space $C^{k,\alpha}(A)$, the following is particularly convenient:

$$\|f\|_{C^{k,\alpha}} := \|\nabla^k f\|_{C^{0,\alpha}_{\mathrm{om}}} + \sum_{h=0}^{k} |\nabla^h f(x_0)| . \tag{3.2}$$

The following interpolation inequality will be useful: if E is a convex set with non-empty interior and f is of class C^1 then[20]

$$\|f\|_{C^{0,\alpha}_{\rm om}} \le 2\,\|f\|_\infty^{1-\alpha}\|\nabla f\|_\infty^\alpha\,. \tag{3.3}$$

3.4. CONSTRUCTION OF C_i, Q_i, g_i, h_i. - For the rest of this section we fix a positive real number α such that

$$\alpha < \frac{1}{d-1}\,. \tag{3.4}$$

We also fix an integer $N > 1$ which is both a d-th and a $(d-1)$-th power of integers, e.g., $N = 2^{d(d-1)}$. Since N is a $(d-1)$-th power of an integer, we can cover the $(d-1)$-dimensional square Q_0 by N squares Q_i with pairwise disjoint interiors and side-length $\rho := N^{-1/(d-1)}$. Since N is the d-th power of an integer, for every positive real number r such that

$$r < N^{-1/d}\,, \tag{3.5}$$

we can find N pairwise disjoint cubes C_i with side-length r contained in the interior of C_0. For every $i = 1, \dots, N$ we set

$$g_i := g_{C_i}\,, \quad h_i := h_{Q_i}$$

(thus g_i and h_i are homotheties with scaling factors r and ρ, respectively).

For the rest of this section we fix r so that it satisfies (3.5) and

$$\rho < r^{1+\alpha}\,. \tag{3.6}$$

This assumption is compatible with (3.5) because the upper bound on α and the definition of ρ imply $\rho < (N^{-1/d})^\alpha$.

Finally, we define the open set

$$A := \mathrm{Int}\big(C_0 \setminus (C_1 \cup \ldots \cup C_N)\big)\,.$$

3.5. CONSTRUCTION OF f_0. - We construct a map $f_0 : \mathbb{R}^d \to Q_0$ with the following properties:

(i) f_0 is of class C^2;

(ii) f_0 agrees with $h_i \circ f_0 \circ g_i^{-1}$ on a neighbourhood of ∂C_i for $i = 1, \dots, N$;

[20] To obtain (3.3) it suffices to estimate the numerator $|f(x) - f(y)|$ at the right-hand side of (3.1) by $2\|f\|_\infty$ when $|x-y| \ge \|f\|_\infty\|\nabla f\|_\infty^{-1}$, and by $\|\nabla f\|_\infty|x-y|$ otherwise.

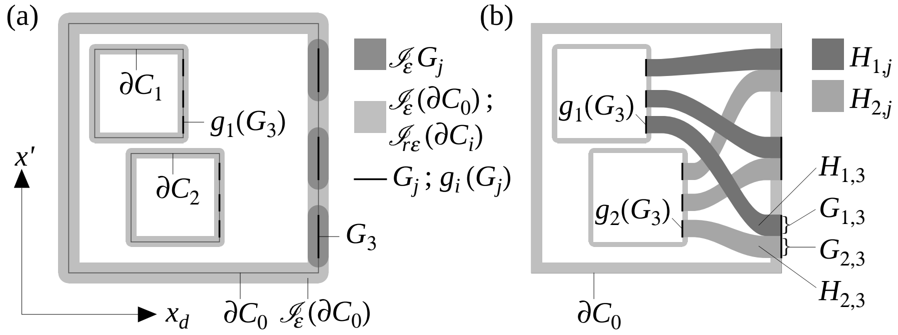

Figure 3. Sets involved in the construction of f_0 on F_1 ($N = 2$): contrary to what appears on the right, in more that two dimensions the sets H_{ij} do not overlap.

(iii) there exists three pairwise disjoint sets G_j contained in ∂C_0 such that the following holds: for every $i = 1, \dots, N$ and every $y \in Q_i$ both $f_0^{-1}(y) \cap G_j$ and $f_0^{-1}(y) \cap g_i(G_j)$ consist of single points, denoted by x_j and x'_j respectively, and there exist pairwise disjoint curves joining x_j and x'_j and contained in $f_0^{-1}(y) \cap A$ except for the end points, cf. Figure 2(a).

The construction of f_0 is divided in three steps.

Step 1. For $j = 1, 2, 3$ we define $G_j := Q'_j \times \{1/2\}$ where Q'_1, Q'_2, Q'_3 are pairwise disjoint squares contained in Q_0. Then we choose $\varepsilon > 0$ so that

a) the ε-neighbourhoods $\mathscr{I}_\varepsilon G_j$ with $j = 1, 2, 3$ are pairwise disjoint;

b) $\mathscr{I}_\varepsilon(\partial C_0)$ and $\mathscr{I}_{r\varepsilon}(\partial C_i)$ are disjoint for $i = 1, \dots, N$ (see Figure 3(a)).

Then we set

$$f_0(x) := h^{-1}_{Q'_j}(x') \quad \text{for every } x \in \mathscr{I}_\varepsilon G_j,$$

and take an arbitrary smooth extension of f_0 to the rest of $\mathscr{I}_\varepsilon(\partial C_0)$.

Step 2. We define f_0 on the sets $\mathscr{I}_{r\varepsilon}(\partial C_i)$ with $i = 1, \dots, N$ so that property (ii) above is satisfied, that is, $f_0 := h_i \circ f_0 \circ g_i^{-1}$.[21]

Step 3. So far the map f_0 has been defined on a neighbourhood of the union of ∂C_i with $i = 0, \dots, N$. Now we extend it to the rest of $\mathbb{R}^d$ so that property (iii) above is satisfied.

Let $y \in Q_i$ be fixed: by Step 1 there exists a unique point in G_j, denoted by $x_j(y)$, such that $f(x_j(y)) = y$, and by Step 2 there exists a unique point $x'_j(y)$ in $g_i(G_j)$ such that $f(x'_j(y)) = y$.

Roughly speaking, the idea is to choose for all $y \in Q_0$ and all $j = 1, 2, 3$ pairwise disjoint curves joining $x_j(y)$ and $x'_j(y)$ and contained in A except the end points, and then set $f_0 := y$ on each curve. However, in order to get a smooth

[21] This definition is well-posed because the sets $\mathscr{I}_\varepsilon(\partial C_0)$ and $\mathscr{I}_{r\varepsilon}(\partial C_i)$ are pairwise disjoint.

map we need a careful construction. For every $j = 1, 2, 3$ and every $i = 1, \dots, N$, the sets

$$G_{ij} := G_j \cap f_0^{-1}(Q_i)$$

are squares with pairwise disjoint interiors of the form $Q''_{ij} \times \{1/2\}$ which cover G_j, while the sets $g_i(G_j)$ are pairwise disjoint squares contained in ∂C_i that can be written as $Q''_{ij} \times \{s_i\}$ for suitable s_i.

Therefore we can find closed sets H_{ij}, with pairwise disjoint interiors and contained in the closure of A, and diffeomorphisms

$$\Psi_{ij} : Q_0 \times [0, 1] \to H_{ij}$$

which map $Q_0 \times \{0\}$ onto $g_i(G_j)$, and $Q_0 \times \{1\}$ onto G_{ij}, see Figure 3(b), and more precisely

$$\Psi_{ij}(x', t) = \begin{cases} (h_{Q''_{ij}}(x'), s_i + t) & \text{for } x' \in Q_0,\ t \in [0, r\varepsilon], \\ \\ (h_{Q'_{ij}}(x'), t - 1/2) & \text{for } x' \in Q_0,\ t \in [1 - \varepsilon, 1]. \end{cases}$$

By Step 1 and Step 2, the map f_0 is already defined on the points of H_{ij} of the form $\Psi_{ij}(x', t)$ with $x' \in Q_0$ and $t \in [0, r\varepsilon] \cup [1 - \varepsilon, 1]$, and satisfies

$$f_0(\Psi_{ij}(x', t)) = h_i^{-1}(x') \,.$$

Then we extend f_0 to the rest of H_{ij} just by setting $f_0(\Psi_{ij}(x', t)) := h_i^{-1}(x')$ for every $x' \in Q_0$ and every $t \in [0, 1]$.

Finally we take an arbitrary extension with compact support of f_0 to the rest of $\mathbb{R}^d$.[22]

3.6. CONSTRUCTION OF f_n. - We denote by J the set of indexes $\{1, \dots, N\}$, and for every $n = 2, 3, \dots$ and every $\mathbf{i} = (i_1, \dots, i_n) \in J^n$ we define

$$\begin{aligned} g_{\mathbf{i}} &:= g_{i_1} \circ g_{i_2} \circ \ldots \circ g_{i_n} \,, \quad C_{\mathbf{i}} := g_{\mathbf{i}}(C_0) \,, \\ h_{\mathbf{i}} &:= h_{i_1} \circ h_{i_2} \circ \ldots \circ h_{i_n} \,, \quad Q_{\mathbf{i}} := h_{\mathbf{i}}(Q_0) \,, \end{aligned} \tag{3.7}$$

and

$$A_n := \mathbb{R}^d \setminus \Big(\bigcup_{\mathbf{i} \in J^n} C_{\mathbf{i}} \Big) \,.$$

For every $n = 1, 2, \dots$ we define the map $f_n : \mathbb{R}^d \to Q_0$ by the recursive relation

$$f_n(x) := \begin{cases} f_{n-1}(x) & \text{if } x \in A_n, \\ \\ h_{\mathbf{i}} \circ f_0 \circ g_{\mathbf{i}}^{-1}(x) & \text{if } x \in C_{\mathbf{i}} \text{ with } \mathbf{i} \in J^n. \end{cases} \tag{3.8}$$

[22] There exists such an extension of class C^2, and even a smooth one, provided that the diffeomorphisms Ψ_{ij} are carefully chosen (we omit the verification).

3.7. PROPOSITION. - *Take α and f_n as in §3.4 and §3.6. Then*

(i) *the maps $f_n : \mathbb{R}^d \to Q_n$ are of class C^2;*

(ii) *the maps f_n converge to a limit map $f : \mathbb{R}^d \to Q_0$ of class $C^{1,\alpha}$;*

(iii) *for every $y \in Q_0$ the level set $E_y := f^{-1}(y)$ contains a triod, and therefore at least one of its connected component is neither a point nor a simple curve.*

PROOF. - We prove statement (i) by induction on n. The map f_0 is of class C^2 by construction (property (i) in §3.5), and if f_{n-1} is of class C^2 then formula (3.8) implies that f_n is of class C^2 provided that f_{n-1} agrees with $h_{\mathbf{i}} \circ f_0 \circ g_{\mathbf{i}}^{-1}$ on a neighbourhood of $\partial C_{\mathbf{i}}$ for every $\mathbf{i} \in J^n$.

For every $\mathbf{i} = (i_1, \dots, i_n) \in J^n$ let $\mathbf{i}' = (i_1, \dots, i_{n-1})$. Formula (3.8) implies that $f_{n-1} = h_{\mathbf{i}'} \circ f_0 \circ g_{\mathbf{i}'}^{-1}$ on $C_{\mathbf{i}'}$, and therefore property (ii) in §3.5 and the definitions of $h_{\mathbf{i}}$ and $g_{\mathbf{i}}$ imply that in a neighbourhood of $\partial C_{\mathbf{i}}$

$$f_{n-1} = h_{i_n} \circ h_{\mathbf{i}'} \circ f_0 \circ g_{\mathbf{i}'}^{-1} \circ g_{i_n}^{-1} .$$

To prove statement (ii) it suffices to show that the maps f_n form a Cauchy sequence with respect to the norm of $C^{1,\alpha}(\mathbb{R}^d)$ defined by (3.2) with $x_0 \in \partial C_0$. Let us estimate the norm of $f_n - f_{n-1}$. By (3.8) this map vanishes outside the union of all $C_{\mathbf{i}}$ with $\mathbf{i} \in J^n$, and then

$$\begin{aligned} \|f_n - f_{n-1}\|_{C^{1,\alpha}(\mathbb{R}^d)} &= \sup_{\mathbf{i} \in J^n} \|\nabla f_n - \nabla f_{n-1}\|_{C^{0,\alpha}_{\mathrm{om}}(C_{\mathbf{i}})} \\ &\leq \sup_{\mathbf{i} \in J^n} \Big[\|\nabla f_n\|_{C^{0,\alpha}_{\mathrm{om}}(C_{\mathbf{i}})} + \|\nabla f_{n-1}\|_{C^{0,\alpha}_{\mathrm{om}}(C_{\mathbf{i}'})} \Big] . \end{aligned} \tag{3.9}$$

By (3.8) the map f_n agrees with $h_{\mathbf{i}} \circ f_0 \circ g_{\mathbf{i}}^{-1}$ on each $C_{\mathbf{i}}$, and since $g_{\mathbf{i}}$ and $h_{\mathbf{i}}$ are homotheties with scaling factors r^n and ρ^n, respectively (cf. (3.7) and §3.4), using the interpolation inequality (3.3) we get

$$\begin{aligned} \|\nabla f_n\|_{C^{0,\alpha}_{\mathrm{om}}(C_{\mathbf{i}})} &\leq 2 \|\nabla f_n\|_\infty^{1-\alpha} \|\nabla^2 f_n\|_\infty^\alpha \\ &= 2 \left(\frac{\rho^n}{r^n} \|\nabla f_0\|_\infty \right)^{1-\alpha} \left(\frac{\rho^n}{r^{2n}} \|\nabla^2 f_0\|_\infty \right)^\alpha \\ &= C \left(\frac{\rho}{r^{1+\alpha}} \right)^n , \end{aligned} \tag{3.10}$$

where $C := 2\|\nabla f_0\|_\infty^{1-\alpha} \|\nabla^2 f_0\|_\infty^\alpha$.[23] Putting together estimates (3.9) and (3.10) we get

$$\|f_n - f_{n-1}\|_{C^{1,\alpha}(\mathbb{R}^d)} \leq 2C \left(\frac{\rho}{r^{1+\alpha}} \right)^{n-1} , \tag{3.11}$$

and therefore assumption (3.6) implies that the sum of $\|f_n - f_{n-1}\|_{C^{1,\alpha}}$ over all $n = 1, 2, \dots$ is finite, which implies that (f_n) is a Cauchy sequence in $C^{1,\alpha}$.

[23] C is finite because f_0 is a map of class C^2 on $\mathbb{R}^d$ with compact support.

We now prove statement (iii). Let $y \in Q_0$ be fixed. We choose a sequence of indexes $i_n \in J$ such that, for every n, the point y belongs to $Q_{\mathbf{i}_n}$ with $\mathbf{i}_n := (i_1, \dots, i_n)$, and we denote by $\bar{x}$ the only point in the intersection of the sets $C_{\mathbf{i}_n}$.

Let $n = 0, 1, \dots$ be fixed. Since f agrees with $h_{\mathbf{i}_n} \circ f_0 \circ g_{\mathbf{i}_n}^{-1}$ on $\partial C_{\mathbf{i}_n}$, property (iii) in §3.5 implies that for every $j = 1, 2, 3$ the set $f^{-1}(y) \cap g_{\mathbf{i}_n}(G_j)$ consists of only one point, denoted by $x_{j,n}$, and there exist pairwise disjoint curves parametrized by $\gamma_{j,n}$ joining $x_{j,n}$ to $x_{j,n+1}$ and contained in $E_y \cap \mathrm{Int}(C_{\mathbf{i}_n} \setminus C_{\mathbf{i}_{n+1}})$ except for the end points.

Upon re-parametrization, we can assume that each $\gamma_{j,n}$ is defined on the interval $[2^{-n-1}, 2^{-n}]$ and satisfies $\gamma_{j,n}(2^{-n}) = x_{j,n}$ and $\gamma_{j,n}(2^{-n-1}) = x_{j,n+1}$. We then set

$$\gamma_j(t) := \begin{cases} \gamma_{j,n}(t) & \text{for } 2^{-n-1} < t \leq 2^{-n},\ n = 0, 1, \dots, \\ \\ \bar{x} & \text{for } t = 0. \end{cases}$$

It is easy to check that the paths γ_j are continuous (the only issue being the continuity at $t = 0$), and the associated curves are contained in the level set E_y and have only the end point $\bar{x}$ in common. Hence the union of these curves is a triod contained in E_y.

A standard topological argument shows that a connected set in $\mathbb{R}^d$ that contains a triod cannot be a simple curve (and clearly not even a point). $\square$

4. Examples of functions without the weak Sard property

The singular set of f a Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}$ is the set S of all $x \in \mathbb{R}^2$ where f is not differentiable or is differentiable and $\nabla f(x) = 0$. As explained in the introduction, we are interested in the following property, which for functions of class C^2 is an immediate corollary of Sard's theorem: the push-forward according to f of the restriction of $\mathscr{L}^2$ to S is singular w.r.t. $\mathscr{L}^1$, that is

$$f_\#(1_S \mathscr{L}^2) \perp \mathscr{L}^1 \,. \tag{4.1}$$

This property clearly implies the following *weak Sard property* (see [1], §****):

$$f_\#(1_{S \cap E^*} \mathscr{L}^2) \perp \mathscr{L}^1 \,, \tag{4.2}$$

where the set E^* is the union of all connected components with positive length of all level sets of f (see §2.4).

As pointed out in [1], Remark ****, property (4.1), and therefore also (4.2), hold whenever f has the Lusin property with functions of class C^2, and in particular when f is locally of class $W^{2,1}$.

In this section we show that in terms of Sobolev classes this regularity assumption is optimal: in §4.1-§4.5 we construct an example of Lipschitz function

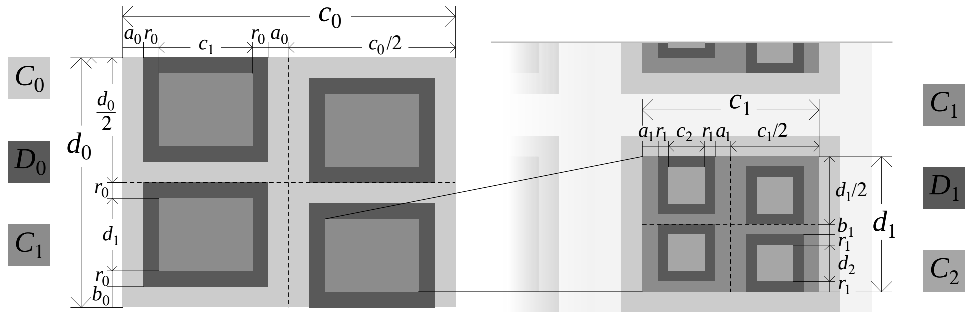

Figure 4. The sets C_i for $i = 0, 1, 2$, and D_i for $i = 0, 1$.

f without the weak Sard property (Proposition 4.7), and in §4.8 we show how to modify this construction so to obtain a function f' of class $C^{1,\alpha}$ for every $\alpha < 1$, and therefore also of class $W^{\beta,p}$ for every $\beta < 2$ and every $p \leq \infty$.

Finally, in Proposition 4.10 we prove that the class $\mathscr{F}$ of all Lipschitz functions $f : \mathbb{R}^2 \to \mathbb{R}$ which satisfy the weak Sard property is residual in the Banach space of Lipschitz functions $\mathrm{Lip}(\mathbb{R}^2)$; this shows that in some sense $\mathrm{Lip}(\mathbb{R}^2)$ is the smallest Banach space of functions that contains $\mathscr{F}$ (Remark 4.11).

4.1. Construction parameters. - For the rest of this section (a_n), (b_n), and (r_n) are decreasing sequences of positive real numbers such that

$$a_n \sim b_n \sim r_n \sim \frac{1}{n^2 2^n}\,, \tag{4.3}$$

where $\sim$ stands for asymptotic equivalence.[24] Hence the following sums are finite:

$$\hat{a} := \sum_{n=0}^{\infty} 2^{n+2} a_n\,, \quad \hat{b} := \sum_{n=0}^{\infty} 2^{n+1} b_n\,, \quad \hat{r} := \sum_{n=0}^{\infty} 2^{n+2} r_n\,.$$

We also choose $\delta > 0$, and set

$$c_0 := \delta + \hat{a} + \hat{r}\,, \quad d_0 := \delta + \hat{b} + \hat{r}\,.$$

4.2. Construction of the sets C_n and D_n. - The set C_0 is a closed rectangle with width c_0 and height d_0; D_0 is the union of 4 closed rectangles (called the *components* of D_0) with width $\frac{1}{2}c_0 - 2a_0$ and height $\frac{1}{2}d_0 - b_0$; C_1 is the union of 4 closed rectangles with width $c_1 := \frac{1}{2}c_0 - 2a_0 - 2r_0$ and height $d_1 := \frac{1}{2}d_0 - b_0 - 2r_0$, each one concentric to a component of D_0. And so on (see Figure 4).

[24] For the purpose of constructing the function f (see §4.5), the parameters r_n could be taken equal to 0; the choice in (4.3) is relevent only in the construction of the function f' (see §4.8).

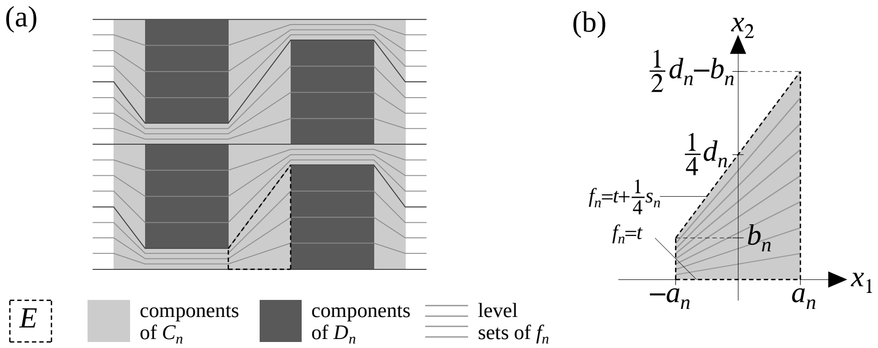

Figure 5. The level curves of f_n in each connected component of C_n

Thus C_n is the union of 4^n pairwise disjoint, closed rectangles with width c_n and height d_n (the components of C_n); D_n is the union of 4^{n+1} pairwise disjoint, closed rectangles with width $\frac{1}{2}c_n - 2a_n$ and height $\frac{1}{2}d_n - b_n$, and each component of C_n contains 4 components of D_n; C_{n+1} is the union of 4^{n+1} pairwise disjoint, closed rectangles with width c_{n+1} and height d_{n+1} given by

$$c_{n+1} := \frac{1}{2}c_n - 2a_n - 2r_n\,, \quad d_{n+1} := \frac{1}{2}d_n - b_n - 2r_n\,, \tag{4.4}$$

and each of these rectangles is concentric to a component of D_n.

Taking into account the definition of c_0 and d_0 it follows immediately that

$$2^n c_n = c_0 - \sum_{m=0}^{n-1} \left(2^{m+2} a_m + 2^{m+2} r_m\right) \searrow \delta\,,$$

and

$$2^n d_n = d_0 - \sum_{m=0}^{n-1} \left(2^{m+1} b_m + 2^{m+2} r_m\right) \searrow \delta\,.$$

In particular c_n and d_n are always strictly positive,[25] and satisfy

$$c_n \sim d_n \sim \frac{\delta}{2^n}\,. \tag{4.5}$$

We denote by C the intersection of the closed sets C_n. Thus (4.5) yields

$$\mathscr{L}^2(C) = \lim_{n\to\infty} \mathscr{L}^2(C_n) = \lim_{n\to\infty} 4^n c_n d_n = \delta^2\,. \tag{4.6}$$

4.3. CONSTRUCTION OF THE FUNCTIONS f_n AND h_n. - We construct a sequence of Lipschitz and piecewise smooth functions $f_n : \mathbb{R}^2 \to \mathbb{R}$ as follows. The function $f_0 = f_0(x_1, x_2)$ agrees with x_2 in $\mathbb{R}^2 \setminus C_0$, while in C_0 it is defined by its level curves, which are described in Figure 5(a); f_1 agrees with f_0 in $\mathbb{R}^2 \setminus C_1$, while in each component of C_1 it is defined by its level curves, described in Figure 5(a). And so on for $n = 2, 3, \ldots$

[25] This means that C_n and D_n are well-defined for every n.

We denote by s_n the oscillation of f_n on the components of C_n.[26] On each component R of D_n the function f_n is affine, depends only on the variable x_2, and is increasing in x_2. Hence $\nabla f_n = (0, v_n)$ for some $v_n > 0$. Clearly the same holds for the components R' of C_{n+1}, and then

$$v_n = \frac{\operatorname{osc}(f_n, R)}{\operatorname{height}(R)} = \frac{\operatorname{osc}(f_n, R')}{\operatorname{height}(R')} .$$

Since $\operatorname{height}(R) = d_{n+1} + 2r_n$, $\operatorname{height}(R') = d_{n+1}$, $\operatorname{osc}(f_n, R) = \frac{1}{4} s_n$ (cf. Figure 5(a)), and $\operatorname{osc}(f_n, R') = s_{n+1}$, the second identity in the previous formula yields

$$4s_{n+1} = \frac{d_{n+1}}{d_{n+1} + 2r_n} s_n ,$$

and taking into account that $s_0 = d_0$ we get

$$4^n s_n = d_0 \left[\prod_{m=1}^{n} \frac{d_m}{d_m + 2r_{m-1}} \right] \searrow d_0 \underbrace{\left[\prod_{m=1}^{\infty} \frac{d_m}{d_m + 2r_{m-1}} \right]}_{\sigma} . \tag{4.7}$$

Using (4.3) and (4.5) we obtain $\log(d_m/(d_m + 2r_{m-1})) = O(-1/m^2)$, and therefore the infinite product σ is strictly positive. In particular, (4.7) implies

$$s_n \sim d_0 \sigma 4^{-n} . \tag{4.8}$$

Finally, for every $n = 1, 2, \dots$ we set

$$h_n := f_n - f_{n-1} . \tag{4.9}$$

4.4. Estimate on the norm of f_n and h_n. - To begin with, we compute the L^∞ norm of ∇f_n on the set C_n. Clearly it suffices to compute the supremum of $|\nabla f_n|$ on one component of C_n, and a closer inspection of Figure 5(a) shows that this supremum is attained in the set E (defined in the same figure). If we choose the axis as in Figure 5(b) and denote by t the value of f_n at the bottom of E, then one readily checks that for every $x \in E$, $f_n(x)$ is given by

$$f_n(x) = t + \frac{a_n s_n x_2}{a_n d_n + (d_n - 4b_n) x_1} ,$$

and a straightforward computation yields

$$\nabla f_n(x) = \frac{(-(d_n - 4b_n)(f_n(x) - t), a_n s_n)}{a_n d_n + (d_n - 4b_n) x_1} . \tag{4.10}$$

[26] This oscillation is clearly the same on all components of C_n.

Equations (4.3) and (4.5) imply $d_n - 4b_n \sim \delta 2^{-n}$; in particular $d_n - 4b_n > 0$ for n sufficiently large, and, taking into account that $x_1 \geq -a_n$, the denominator in (4.10) is larger than $4a_nb_n$. Since $|f_n - t| \leq s_n$, the absolute value of the first component of the numerator is smaller than $(d_n - 4b_n)s_n$. Hence (for n sufficiently large)

$$\|\nabla f_n\|_{L^\infty(C_n)} \leq \frac{s_n}{4b_n}\left(\frac{d_n - 4b_n}{a_n} + 1\right) = O(n^4 2^{-n}), \tag{4.11}$$

where the final equality follows from equations (4.3) and (4.8).

Now we focus on the functions h_n defined in (4.9). By construction, f_n and f_{n-1} agree outside C_n, and therefore the support of h_n is contained in C_n. Then (4.11) yields

$$\|\nabla h_n\|_\infty \leq \|\nabla f_n\|_{L^\infty(C_n)} + \|\nabla f_{n-1}\|_{L^\infty(C_{n-1})} = O(n^4 2^{-n}), \tag{4.12}$$

and since the distance of a point in C_n from $\mathbb{R}^2 \setminus C_n$ is of order $c_n = O(2^{-n})$,

$$\|h_n\|_\infty = O(n^4 4^{-n}). \tag{4.13}$$

4.5. Construction of f. - We take the functions f_n and h_n as in §4.3, and for every $x \in \mathbb{R}^2$ we set

$$f(x) := \lim_{n\to+\infty} f_n(x) = f_0(x) + \sum_{n=1}^{\infty} h_n(x). \tag{4.14}$$

Estimates (4.12) and (4.13) show that the sum of the norms $\|h_n\|_{C^{0,1}}$ (defined by (3.2) with an arbitrary choice of x_0) is finite and therefore f is a well-defined Lipschitz function on $\mathbb{R}^2$.

4.6. Lemma. - *Let be given a compact metric space X and a sequence of countable Borel partitions $\mathscr{F}_n$ of X. Let δ_n be the supremum of the diameters of the elements of $\mathscr{F}_n$, and assume that $\delta_n \to 0$ as $n \to +\infty$. Then two finite measures μ_1 and μ_2 on X which agree on the elements of $\cup\mathscr{F}_n$ agree also on all Borel sets.*

Proof. - For every n, let $\mathscr{G}_n$ be the class of all bounded functions on X which are constant on each element of $\mathscr{F}_n$. Then $\int g\, d\mu_1 = \int g\, d\mu_2$ for every $g \in \mathscr{G}_n$, and the equality carries over to all continuous functions g because they can be uniformly approximated by functions in $\cup\mathscr{G}_n$. □

4.7. Proposition. - *Take the set C as in §4.2, and let f be the function defined in* (4.14). *Then*

(i) *f is differentiable at every $x \in C$ and $\nabla f(x) = 0$;*

(ii) $\mathscr{L}^1(f(C)) = d_0\sigma$ *where* σ *is given in* (4.7)*;*

(iii) $f_\#(1_C \cdot \mathscr{L}^2) = m 1_{f(C)} \cdot \mathscr{L}^1$ *where* $m := \delta^2/(d_0\sigma)$*; in particular,* f *does not satisfy the weak Sard property* (4.2).

Proof. - (i) For every $n \geq 0$, we write f as $f_n + h_{n+1} + h_{n+2} + \cdots$, and then estimates (4.11) and (4.12) yield

$$\|\nabla f\|_{L^\infty(C_n)} = O(n^4 2^{-n}) \, .$$

This implies that the Lipschitz constant of f on each component of C_n is of order $O(n^4 2^{-n})$; since C is contained in the interior of C_n, it follows that

$$\limsup_{y \to x} \frac{|f(y) - f(x)|}{|x - y|} = O(n^4 2^{-n})$$

for every $x \in C$, and taking the limit as $n \to +\infty$ we obtain that f is differentiable at x and $\nabla f(x) = 0$.

(ii) We note that $f(C)$ is the intersection of all $f(C_n)$, and $f(C_n)$ agrees with $f_n(C_n)$, and therefore it is the union of 4^n pairwise disjoint, closed interval with length s_n – the images according to f of the components of C_n. Therefore, using (4.8) we get

$$\mathscr{L}^1(f(C)) = \lim_{n \to +\infty} \mathscr{L}^1(f_n(C_n)) = \lim_{n \to +\infty} 4^n s_n = d_0\sigma \, .$$

(iii) We must show that the measures $\mu := f_\#(1_C \cdot \mathscr{L}^2)$ and $\lambda := m 1_{f(C)} \cdot \mathscr{L}^1$ are the same. Since both μ and λ are supported on the compact set $f(C)$, we apply Lemma 4.6 to the partitions $\mathscr{F}_n$ given by the sets $R' := f(R \cap C)$ where R is a component of C_n, and deduce that it suffices to prove $\mu(R') = \lambda(R')$ for every such R'.

Since C can be written as a disjoint union of 4^n translated copies of $R \cap C$, we have

$$\mu(R') = \mathscr{L}^2(R \cap C) = 4^{-n} \mathscr{L}^2(C) = 4^{-n}\delta^2 \, .$$

On the other hand $f(C)$ can be written as a disjoint union of 4^n translated copies of R', and then

$$\lambda(R') = m\,\mathscr{L}^1(R') = 4^{-n} m\,\mathscr{L}^1(f(C)) = 4^{-n} m d_0\sigma = 4^{-n}\delta^2 \, . \qquad \square$$

In the next paragraphs we show how to modify the construction in §4.5 so to obtain a more regular example.

4.8. Construction of f'. - For every $n = 0, 1, 2, \ldots$ we consider the smoothing kernels

$$\rho_n(x) := r_n^{-2} \rho(x/r_n) \, ,$$

where r_n is given in §4.1 and $\rho : \mathbb{R}^2 \to \mathbb{R}$ is a positive smooth function which satisfies the following conditions:

(a) the support of ρ is contained in the ball with center 0 and radius $1/2$;

(b) $\int_{\mathbb{R}^2} \rho(x)\, dx = 1$ and $\int_{\mathbb{R}^2} x\, \rho(x)\, dx = 0$.

Then we take f_n and h_n as in §4.3, and define $f' : \mathbb{R}^2 \to \mathbb{R}$ as

$$f' := f_0 * \rho_0 + \sum_{n=1}^{\infty} h_n * \rho_n\,, \tag{4.15}$$

where $*$ denotes the usual convolution product.

4.9. Proposition. - *Take C and f as in §4.2 and §4.5, and let f' be the function defined by* (4.15)*. Then*

(i) *f' is a well-defined function of class $C^{1,\alpha}$ for every $\alpha < 1$;*

(ii) *f' is smooth on $\mathbb{R}^2 \setminus C$ and agrees with f on C;*

(iii) *$\nabla f'(x) = 0$ for every $x \in C$;*

(iv) *$f'_\#(1_C \cdot \mathscr{L}^2) = m 1_{f'(C)} \cdot \mathscr{L}^1$ where $m := \delta^2/(d_0\sigma)$; in particular, f' does not satisfy the weak Sard property* (4.2)*.*

Proof. - (i) We claim that the series $\sum h_n * \rho_n$ converge in norm in the Banach space $C^{1,\alpha}$ for every $\alpha < 1$, and therefore f' belongs to this space. Indeed, using (3.3), (4.12), and (4.3) we obtain

$$\begin{aligned}
\|\nabla(h_n * \rho_n)\|_{C^{0,\alpha}_{\mathrm{om}}} &\leq 2\|\nabla(h_n * \rho_n)\|_\infty^{1-\alpha} \|\nabla^2(h_n * \rho_n)\|_\infty^{\alpha} \\
&\leq 2\big(\|\nabla h_n\|_\infty \|\rho_n\|_1\big)^{1-\alpha} \big(\|\nabla h_n\|_\infty \|\nabla \rho_n\|_1\big)^{\alpha} \\
&= 2\|\nabla h_n\|_\infty \|\nabla\rho_n\|_1^{\alpha} \\
&= \|\nabla h_n\|_\infty O(r_n^{-\alpha}) = O(n^{4+2\alpha} 2^{-(1-\alpha)n})\,.
\end{aligned}$$

Moreover estimates (4.12) and (4.13) hold even if h_n is replaced by $h_n * \rho_n$, and therefore $\|h_n * \rho_n\|_{C^{1,\alpha}} = O(n^6 2^{-(1-\alpha)n})$, which implies that the sum of the norms $\|h_n * \rho_n\|_{C^{1,\alpha}}$ is finite.

(ii) Each h_n is supported in C_n (cf. §4.3) and since ρ_n is supported in the ball with center 0 and radius $r_n/2$, the function $h_n * \rho_n$ is supported in the closed set C'_n of all points whose distance from C_n is at most $r_n/2$. Since the sets C'_n decrease to C, f' agrees with a finite sum of smooth functions in every open set with positive distance from C, and therefore is smooth on $\mathbb{R}^2 \setminus C$.

By construction, f_n is affine on D_n and then $h_n = f_n - f_{n-1}$ is affine on D_n too. Therefore, assumption (b) in §4.8 and the fact that ρ_n is supported in the ball with center 0 and radius $r_n/2$ imply that $h_n * \rho_n = h_n$ in the closed set D'_n of all points of D_n whose distance from the boundary of D_n is at least $r_n/2$. Since D'_n contains C_{n+1} (cf. Figure 4), which in turn contains C, we have that

$h_n * \rho_n = h_n$ on C. A similar argument shows that $f_0 * \rho_0 = f_0$ on C. Then $f = f'$ on C by definitions (4.14) and (4.15).

Since $f' = f$ on C, statements (iii) and (iv) follow (almost) immediately from statements (i) and (iii) of Proposition 4.7. $\square$

We conclude with a result on the genericness of the weak Sard property.

4.10. PROPOSITION. - *Let* $\mathrm{Lip}(\mathbb{R}^2)$ *be the Banach space of Lipschitz function on* $\mathbb{R}^2$*, and let* $\mathscr{G}$ *be the subclass of all functions whose singular set has measure zero. Then* $\mathscr{G}$ *is residual in* $\mathrm{Lip}(\mathbb{R}^2)$.[27]

PROOF. - For every $r > 0$ let $\mathscr{G}_r$ be the class of all $f \in \mathrm{Lip}(\mathbb{R}^2)$ whose singular set $S(f)$ satisfies $\mathscr{L}^2(S(f)) < r$. Since $\mathscr{G}$ is the intersection of all $\mathscr{G}_r$ with $r = 1/n$ and $n = 1, 2, \dots$, it suffices to show that each $\mathscr{G}_r$ is open and dense in $\mathrm{Lip}(\mathbb{R}^2)$. Among the many (equivalent) complete norms on $\mathrm{Lip}(\mathbb{R}^2)$ we consider the following: $\|f\| := |f(0)| + \|\nabla f\|_\infty$.

Step 1: $\mathscr{G}_r$ is open. Take $f \in \mathscr{G}_r$. Since $\mathscr{L}^2(S(f)) < r$, there exists $\delta > 0$ such that $\mathscr{L}^2(\{x : |\nabla f(x)| \le \delta\}) < r$, and therefore $\mathscr{L}^2(S(g)) < r$ for every $g \in \mathrm{Lip}(\mathbb{R}^2)$ such that $\|\nabla f - \nabla g\|_\infty < \delta$, that is, for every g in a neighbourhood of f.

Step 2: $\mathscr{G}_r$ is dense. Given $f \in \mathrm{Lip}(\mathbb{R}^2)$ and $\delta > 0$, we must find $g \in \mathscr{G}_r$ such that $\|g - f\| \le \delta$. We take an open set A which contains $S(f)$ and satisfies $\mathscr{L}^2(A \setminus S(f)) < r$, and choose a sequence of pairwise disjoint closed discs D_n contained in A which cover almost all of $S(f)$. Then for every n we choose a Lipschitz function g_n such that $\nabla g_n \neq 0$ a.e. in D_n and $\nabla g_n = 0$ a.e. in $\mathbb{R}^2 \setminus D_n$. Up to rescaling we can assume that $\|g_n\| \le 2^{-n}\delta$, and then we set

$$g := f + \sum_{n=1}^{\infty} g_n \ .$$

Thus g is Lipschitz, $\|g - f\| \le \delta$, and $S(g)$ is contained in $A \setminus S(f)$, and in particular $\mathscr{L}^2(S(g)) < r$, that is, g belongs to $\mathscr{G}_r$. $\square$

4.11. REMARK. - The class $\mathscr{F}$ of all functions in $\mathrm{Lip}(\mathbb{R}^2)$ which satisfy the weak Sard property contains $\mathscr{G}$, and therefore Proposition 4.10 implies that $\mathscr{F}$ is residual in $\mathrm{Lip}(\mathbb{R}^2)$. It follows that a Banach space X which embeds continuously in $\mathrm{Lip}(\mathbb{R}^2)$ and (whose image) contains $\mathscr{F}$ must agree with $\mathrm{Lip}(\mathbb{R}^2)$ itself: indeed X would be residual in $\mathrm{Lip}(\mathbb{R}^2)$, and then the Open Mapping Theorem implies that the embedding $i : X \to \mathrm{Lip}(\mathbb{R}^2)$ is open, and therefore is an isomorphism.

[27] A set in a topological space is residual if it contains a countable intersection of open dense sets. By Baire theorem, a residual set in a complete metric space is dense.

5. Non-locality of the divergence operator

In this section we show that the divergence operator is not strongly local on the space of vector fields in L^∞ with distributional divergence in L^∞. More precisely, in §5.4 we construct a bounded vector field b on the plane whose distributional divergence $\operatorname{div} b$ belongs to L^∞, is non-trivial, and its support is contained in the set where b takes the value 0 (Proposition 5.6).

Starting from this example, in §5.7 we construct another Lipschitz function f on the plane without the weak Sard property.

5.1. Remark. - There is actually a deep relation between the non-locality of $\operatorname{div} b$ and the lack of weak Sard property of f: in our example b can be written as $b = a\,(\nabla f)^\perp$ where a is the characteristic function of an open set (Proposition 5.8),[28] and it is proved in [1], ****, that if a vector field b in L^∞ with divergence in L^∞ can be represented as $b = a\,(\nabla f)^\perp$ where a is a function in L^∞ and f a Lipschitz function with the weak Sard property, then the divergence of b is local, in the sense that $\operatorname{div} b = 0$ a.e. on every set where b is a.e. constant.

5.2. Construction parameters. - For the rest of this section (d_n) is a fixed, decreasing sequence of positive real numbers such that

$$d_n \geq 5 \cdot 2^{-n} \quad \text{for } n = 0, 1, 2, \dots \tag{5.1}$$

and the sums

$$c_0 := \sum_{n=0}^{\infty} 2^n d_{2n}\,, \quad c_1 := \sum_{n=0}^{\infty} 2^n d_{2n+1} \tag{5.2}$$

are finite. We then choose $r > 0$, and set $a_0 := c_0 + \sqrt{2}r$, $a_1 := c_1 + r$.

5.3. Construction of the sets C and E. - The set C_0 is a closed rectangle with width a_1 and height a_0; C_1 is the union of two closed rectangles obtained by removing from C_0 the middle rectangle with same width (that is, a_1) and height d_0; C_2 is the union of four closed rectangles obtained by removing from each of the two rectangles that compose C_1 the middle rectangle with same height (denoted by a_2) and width d_1. And so on (see Figure 6).

Thus C_n is contained in C_{n-1}, and consists of 2^n closed rectangles with the same width and the same height, called *components* of C_n. Let a_n denote the width of such rectangles for n odd, and the height for n even (thus the length of the other side is a_{n+1}); then a_n satisfies the recursive relation

$$2a_{n+2} = a_n - d_n\,. \tag{5.3}$$

[28] Through this section we write $v^\perp$ and $v^\top$ for the counter-clockwise and clockwise by 90° of the vector $v \in \mathbb{R}^2$.

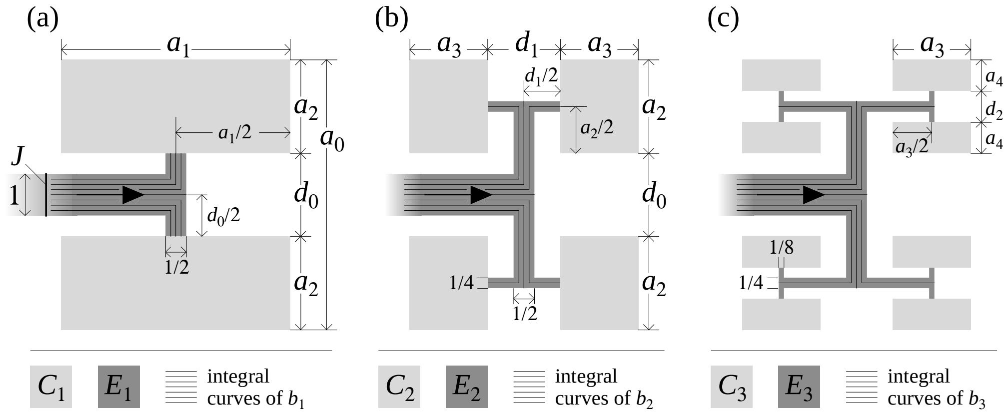

Figure 6. The sets C_i, E_i, and the integral curves of b_i (within E_i) for $i = 1, 2, 3$.

We easily deduce from (5.3) that for every $n \geq 1$

$$2^n a_{2n} = a_0 - \sum_{m=0}^{n-1} 2^m d_{2m} \searrow \sqrt{2}r\,,$$

and

$$2^n a_{2n+1} = a_1 - \sum_{m=0}^{n-1} 2^m d_{2m+1} \searrow r\,.$$

Hence a_n is strictly positive and

$$a_n \sim 2^{(1-n)/2} r\,. \tag{5.4}$$

The set E_n (defined in Figure 6) is open, contained in the complement of C_n, and contains E_{n-1}. Note that E_1 – and therefore all E_n – contains an horizontal strip unbounded to the left.

Finally, we denote by C the intersection of the closed sets C_n, and by E the union of the open sets E_n. Thus C is the product of two Cantor-like sets, and (5.4) yields

$$\mathscr{L}^2(C) = \lim_{n\to\infty} \mathscr{L}^2(C_n) = \lim_{n\to\infty} 2^n a_n a_{n+1} = \sqrt{2}\, r^2\,. \tag{5.5}$$

5.4. Construction of the vector field b. - For every $n \geq 0$, the vector field b_n is set equal to 0 in $\mathbb{R}^2 \setminus E_n$, while in E_n it is piecewise constant, satisfies $|b_n| = 1$ everywhere, and its direction can be inferred from the integral curves and the arrows drawn in Figure 6.[29] Note that b_n agrees with b_{n-1} in E_{n-1} for every n.

[29] In what follows it does not matter how b_n is defined at the corners of the integral curves.

Finally, the vector field b is the pointwise limit of b_n.[30]

The result in the next paragraph is used to compute the divergence of b.[31]

5.5. Divergence of piecewise regular vector fields. - Let b be a bounded vector field on $\mathbb{R}^d$, and $\{A_i\}$ a finite family of pairwise disjoint open sets with Lipschitz boundaries whose closures cover $\mathbb{R}^d$. We assume that b agrees on each A_i with a vector field b_i of class C^1 defined on $\mathbb{R}^d$. Then the distributional divergence of b is a measure of the form

$$\operatorname{div} b = f \cdot \mathscr{L}^d + g\, 1_S \cdot \mathscr{H}^{d-1}\,, \tag{5.6}$$

where S is the union of all ∂A_i, f is the pointwise divergence of b (defined for all $x \notin S$), and g is the difference of the normal components of b at the two sides of S. More precisely, for every $x \in \partial A_i \cap \partial A_j$ where both ∂A_i and ∂A_j admit a tangent plane T, we set $g(x) := (b_i(x) - b_j(x)) \cdot \eta$ where η is the unit vector normal to T that points toward A_i.

5.6. Proposition. - *Take C and b as in §5.3 and §5.4. Then $b = 0$ on C, and $\operatorname{div} b$ is non trivial, belongs to $L^\infty(\mathbb{R}^2)$, and is supported on C. More precisely $\operatorname{div} b = m\, 1_C$ where $m := 1/(\sqrt{2}r^2)$.*

Proof. - The fact that $b = 0$ on C follows immediately from the fact that $b_n = 0$ on C by construction and b is the pointwise limit of b_n.

Step 1. Computation of the divergence of b_n. We compute $\operatorname{div} b_n$ using formula (5.6). In this case $f = 0$ because b_n is piecewise constant, and S consists of the boundary of E_n plus the discontinuity points of b_n contained in E_n, that is, the corner points of the integral curves drawn in Figure 6. Moreover $g = 1$ in $F_n := \partial E_n \cap \partial C_n$, and $g = 0$ everywhere else. Thus $\operatorname{div} b_n = \mu_n$, where μ_n is the restriction of $\mathscr{H}^1$ to the set F_n. Since F_n is the union of 2^n segments of length 2^{-n}, μ_n is a probability measure.

Step 2. Computation of the divergence of b. Since b_n converge to b in L^1, the divergence of b is the limit (in the sense of distributions) of $\operatorname{div} b_n = \mu_n$ and therefore is a probability measure, which we denote by μ. Since the support of each μ_n is contained in C_n, the support of μ is contained in C.

Take $n = 1, 2, \dots$, and let R be one of the 2^n components of C_n. We claim that

$$\mu(R \cap C) = \mu(R) = 2^{-n}\,. \tag{5.7}$$

Indeed, for every $m \geq n$, the rectangle R contains 2^{m-n} of the 2^m segments that compose F_m. Hence $\mu_m(R) = 2^{-n}$, and passing to the limit as $m \to +\infty$

[30] By construction the sequence $b_n(x)$ is definitely constant for every $x \in \mathbb{R}^2$, and therefore it converges to some limit $b(x)$. It can be easily shown that it is a Cauchy sequence in $L^1(\mathbb{R}^2)$.

[31] Through this section, divergence, gradient, and curl are intended in the sense of distributions.

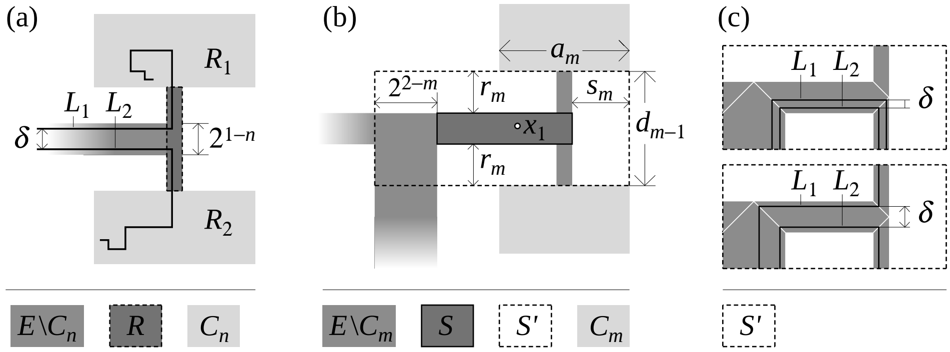

Figure 7. Estimating the distance between the level sets L_1 and L_2

we obtain $\mu(R) \geq 2^{-n}$. On the other hand, if R' is an open neighbourhood of R which intersect no other component of C_n we have $\mu_m(R') = 2^{-n}$ for every $m \geq n$, and passing to the limit we get $\mu(R) \leq \mu(R') \leq 2^{-n}$. Thus (5.7) is proved.

We must show that $\mu = \lambda$ where $\lambda := m\, 1_C \cdot \mathscr{L}^2$. Since μ and λ are supported on C, by Lemma 4.6 it suffices to show that $\mu(R \cap C) = \lambda(R \cap C)$ for every rectangle R taken as in Step 2.

Since C can be written as disjoint union of 2^n translated copies of $R \cap C$, then

$$\lambda(R \cap C) = m\, \mathscr{L}^2(R \cap C) = 2^{-n} m\, \mathscr{L}^2(C) = 2^{-n} = \mu(R \cap C)\,,$$

where the third identity follows from (5.5), and the fourth one from (5.7). □

5.7. Construction of f. - *Step 1. Construction of f on E.* By Proposition 5.6, the divergence of b vanishes on the open set E, and therefore the rotated vector field $b^\top$ (see footnote 28) is curl-free on E. Since the set E is simply connected, $b^\top$ is the gradient of a function $f : E \to \mathbb{R}$ of class $W^{1,\infty}$, and therefore we can assume that f is locally Lipschitz on E.

Clearly $b = (\nabla f)^\perp$ a.e. on E, and the level sets of f agree (in E) with the integral curves of b.

Step 2. f has Lipschitz constant 1 *on E.* Since $|\nabla f| = |b| = 1$ a.e. on E, the restriction of f to every convex subset of E has Lipschitz constant 1, but proving that f has Lipschitz constant 1 on E is more delicate.

Given $t_i \in f(E)$ for $i = 1, 2$, let $L_i := \{x \in E : f(x) = t_i\}$ be the corresponding level sets, and set $\delta := |t_1 - t_2|$. Then f has Lipschitz constant 1 if for every choice of t_i and of $x_i \in L_i$ there holds

$$|x_1 - x_2| \geq \delta\,. \tag{5.8}$$

By a density argument, it suffices to prove (5.8) when both level sets intersect C_n for all n.

Let n be the smallest integer such that $L_1 \cap C_n$ and $L_2 \cap C_n$ are contained in different components of C_n. Denote these components by R_1 and R_2, respectively, and take the rectangle R as in Figure 7(a). This picture clearly shows that if both x_1 and x_2 belong to $R_1 \cup R_2 \cup R$ then (5.8) holds.

We assume now that $x_1 \notin R_1 \cup R_2 \cup R$ (the case $x_2 \notin R_1 \cup R_2 \cup R$ is equivalent). Take m such that x_1 belongs to the rectangle S in Figure 7(b). There are two possibilities: either x_2 belongs to the rectangle S' defined in Figure 7(b), or it doesn't. In the former case, Figure 7(c) shows that (5.8) holds.[32] In the latter case, Figure 7(b) shows that (5.8) holds provided that the lengths r_m, s_m (defined in the figure) and 2^{2-m} are larger than δ.

Let us prove these inequalities. Since $m \le n$ and $2^{1-n} \ge \delta$ (see Figure 7(a)), then

$$2^{2-m} \ge 2^{2-n} \ge 2\delta \,.$$

Moreover assumption (5.1) yields

$$r_m = \frac{1}{2}(d_{m-1} - 2^{1-m}) \ge 2^{2-m} \ge 2\delta \,.$$

Finally, (5.3) and the positivity of a_n imply $a_n \ge d_n$, and taking into account (5.1)

$$s_m = \frac{1}{2}(a_m - 2^{-m}) \ge \frac{1}{2}(d_m - 2^{-m}) \ge 2^{1-m} \ge \delta \,.$$

Step 3. Definition of f outside E. We use McShane lemma to extend f to the rest of $\mathbb{R}^2$ so that the Lipschitz constant is equal to 1.

5.8. Proposition. - *Take C, E, b and f as in the previous paragraphs. Then*
(i) $b = 1_E\,(\nabla f)^\perp$ *a.e.;*
(ii) $\nabla f = 0$ *a.e. in C;*
(iii) f *does not satisfy the weak Sard property.*

Proof. - Statement (i) follows immediately from the definitions of b and f.

(ii). Let $\{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$. Fix a point $x \in C$ where f is differentiable, and for every $n = 1, 2, \dots$ denote by R_n the connected component of C_n which contains x.

Note that R_n has sides of length a_n and a_{n+1} and since $a_n \ge a_{n+1}$, at least one of the points $x \pm \frac{1}{2} a_{n+1} e_1$ belongs to R_n. We denote this point by x_n. Then (5.4) yields

$$|f(x_n) - f(x)| \sim |D_1 f(x)|\, \frac{a_{n+1}}{2} \sim |D_1 f(x)|\, \frac{r}{2} 2^{-n/2} \,. \tag{5.9}$$

On the other hand, Figure 6 shows that the integral curves of b (in E) which intersect R_n are exactly those which intersect a certain (open) segment J_n of

[32] The upper picture in Figure 7(c) refers to the case $m < n$ and the lower one to the case $m = n$.

length 2^{-n} contained in the segment J in Figure 6(a). Taking into account that these integral curves are the restriction to E of the level sets of f, we deduce that $f(C \cap R_n) = f(J_n)$, and since the slope of f along J_n is 1, the set $f(J_n)$ is an interval of length 2^{-n}, that is

$$\operatorname{osc}(f, C \cap R_n) = \operatorname{osc}(f, J_n) = 2^{-n}\,, \tag{5.10}$$

where $\operatorname{osc}(f, E)$ stands for the oscillation of f on the set E.

A close examination of Figure 6 shows that the distance of any point of R_n from $C \cap R_n$ is at most $\frac{1}{2}(d_n + d_{n+1})$. Moreover $d_n = o(2^{-n/2})$ because we assumed that the sums in (5.2) converge. Therefore (5.10) and the fact that the Lipschitz constant of f is 1 yields

$$\operatorname{osc}(f, R_n) \leq \operatorname{osc}(f, C \cap R_n) + d_n + d_{n+1} = o(2^{-n/2})\,. \tag{5.11}$$

Now, (5.9) and (5.11) can be both true only if $D_1 f(x) = 0$. A similar argument shows that $D_2 f(x) = 0$. We conclude that $\nabla f(x) = 0$ for every $x \in C$ where f is differentiable, that is, for a.e. $x \in C$.

(iii). As pointed out before, $f(C) = f(J)$ is an interval of length 1, and therefore, upon addition of a constant, we can assume that this interval is $[0, 1]$. To prove statement (iii) we show that $\mu := f_\#(1_C \cdot \mathscr{L}^2)$ agrees with $m\, 1_{[0,1]} \cdot \mathscr{L}^1$ where $m := \mathscr{L}^2(C)$.

To this end, it suffices to show that μ is non atomic and $\mu(I_{n,k}) = m/2^n$ for every interval $I_{n,k} := ((k-1)2^{-n}, k2^{-n})$ with $n = 1, 2, \dots$ and $k = 1, \dots, 2^n$ (cf. Lemma 4.6).

The latter claim follows by the fact that $C \cap f^{-1}(I_{n,k})$ agrees up to a null set with $C \cap R_{n,k}$ where $R_{n,k}$ is a suitable connected component of C_n, and therefore $\mathscr{L}^2(C \cap R_{n,k}) = m2^{-n}$. The former claim follows by a similar argument. □

6. Appendix: A measurability lemma

Let be given a compact metric space X, a topological space Y, a continuous map $f : X \to Y$, and a real number $d \geq 0$. For every $y \in Y$ we write E_y for the level set $f^{-1}(y)$, E_y^d for union of all connected components C of E_y such that $\mathscr{H}^d(C) > 0$, and E^d the union of E_y^d over all $y \in Y$.

6.1. Proposition. - *The sets E_y^d and E^d are countable unions of closed sets in X; in particular they are Borel measurable.*

In order to prove this statement we need to recall some definitions and known facts. As usual $\mathscr{F}(X)$ is the class of all non-empty closed subsets of X, and is endowed with the Hausdorff distance d_H (see (2.1)). Since X is compact, the metric space $\mathscr{F}(X)$ is compact, and the subclass $\mathscr{F}_c(X)$ of all elements of $\mathscr{F}(X)$ which are *connected* is closed (see for instance [8], Theorems 3.16 and 3.18).

6.2. Lemma. - *If $\mathscr{G}$ is a closed subclass of $\mathscr{F}(X)$, then the union C of all sets in $\mathscr{G}$ is a closed subset of X.*

Proof. - Let x be the limit of a sequence of points $x_n \in C$. For every n there exists $C_n \in \mathscr{G}$ such that $x_n \in C_n$, and since $\mathscr{G}$ is closed in $\mathscr{F}(X)$ and therefore compact, the sequence (C_n) has an accumulation point $C_\infty \in \mathscr{G}$. Hence x belongs to C_∞, and therefore also to C. □

For every set E in X the d-dimensional Hausdorff measure $\mathscr{H}^d(E)$ is defined (up to a renormalization factor) as the supremum over all $\delta > 0$ of the outer measures $\mathscr{H}^d_\delta(E)$ given by

$$\mathscr{H}^d_\delta(E) := \inf \left\{ \sum_i \big(\mathrm{diam}(A_i)\big)^d \right\}, \tag{6.1}$$

where the infimum is taken over all countable coverings $\{A_i\}$ of E such that $\mathrm{diam}(A_i) \leq \delta$ for every i. Note that the value of $\mathscr{H}^d_\delta(E)$ does not change if the sets A_i are assumed to be open.

6.3. Lemma. - *The outer measures $\mathscr{H}^d_\delta$ are upper semicontinuous on $\mathscr{F}(X)$.*

Proof. - Take $E_0 \in \mathscr{F}(X)$ and $\varepsilon > 0$. By (6.1) there exists an open covering $\{A_i\}$ of E_0 such that $\mathrm{diam}(A_i) \leq \delta$ for every i and

$$\sum_i \big(\mathrm{diam}(A_i)\big)^d < \mathscr{H}^d_\delta(E_0) + \varepsilon \, .$$

Since the union A of all A_i is open and the set E_0 is compact, then A contains all sets E in some neighbourhood of E_0, and therefore $\mathscr{H}^d_\delta(E) \leq \mathscr{H}^d_\delta(E_0) + \varepsilon$ for all such E. □

Proof of Proposition 6.1. - We only prove the result for E^d; the proof for E^d_y is essentially the same (and actually slightly simpler).

Let $\mathscr{G}$ be the class of all connected, closed sets C in X which are contained in E_y for some $y \in Y$. It is easy to show that $\mathscr{G}$ is closed in $\mathscr{F}(X)$, and therefore Lemma 6.3 yields that the subclass $\mathscr{G}(\delta, \varepsilon)$ of all $C \in \mathscr{G}$ such that $\mathscr{H}^d_\delta(C) \geq \varepsilon$ is also closed in $\mathscr{F}(X)$ for every $\delta, \varepsilon > 0$. Hence the union $G(\delta, \varepsilon)$ of all sets C in $\mathscr{G}_c(\delta, \varepsilon)$ is closed in X by Lemma 6.2.

To conclude the proof it suffices to verify that E^d agrees with the union of all $G(1/m, 1/n)$ with $m, n = 1, 2, \dots$ □

7. Appendix: Approximation of Lipschitz functions

In this section we deal with the following problem: given a set X endowed with two different distances d and d', is it possible to approximate a real-valued function f on X which is Lipschitz with respect to d by a sequence of functions which are uniformly Lipschitz with respect to d and Lipschitz with respect to d'? The answer is clearly positive if $d' \geq s\,d$ for some constant $s > 0$, but in general it may be negative, even when d and d' induce the same topology.

In Proposition 7.8 we give a necessary and sufficient condition (on d and d') for the existence of such approximation. In Proposition 7.10 we show that such condition is satisfied in a particular case that occurs in the proof of Lemma 2.14 (and is also used in [1]).

7.1. NOTATION. - We recall here some notation specific of this section.

The supremum norm of a real-valued function f is denoted by $\|f\|$.

By *distance* on the set X we mean a function $d : X \times X \to [0, +\infty)$ which satisfies the usual properties; if d satisfies all properties except the implication $d(x, y) = 0 \Rightarrow x = y$, then it is called a *semi-distance.* Two distances d, d' on the set X are topologically equivalent if they induce the same topology on X. They are Lipschitz equivalent if there exist constants $0 < c_1 \leq c_2 < +\infty$ such that $c_1 d \leq d' \leq c_2 d$.

Let d_1, d_2 be distances on X_1, X_2, respectively; for every couple of points $x = (x_1, x_2)$, $y = (y_1, y_2)$ in $X_1 \times X_2$ we define the product distance[33]

$$[d_1 \times d_2](x, y) := d_1(x_1, y_1) + d_2(x_2, y_2)\,.$$

Given a real-valued function f defined on a subset of X, the Lipschitz constant of f with respect to the distance d is denoted by $L(f, d)$ – the value being taken equal to $+\infty$ if f is not Lipschitz. We denote by $\mathrm{Lip}(X, d)$ the class of all Lipschitz functions defined on X. Note that these definitions makes sense even if d is a semi-distance.

7.2. INVOLUTION OF DISTANCES. - Let d_1, d_2 be semi-distances on X. The *involution* of d_1 and d_2 is defined by

$$[d_1 \wedge d_2](x, y) := \inf \left\{ \sum_{i=1}^{n} \min\{d_1(z_{i-1}, z_i), d_2(z_{i-1}, z_i)\} \right\}$$

for every $x, y \in X$, where the infimum is taken over all n and all finite sequences $z_0, \dots, z_n \in X$ that satisfy $z_0 = x$ and $z_n = y$. One readily checks that $d_1 \wedge d_2$ is a semi-distance; indeed, it is the largest among all semi-distances d that satisfy $d \leq d_1, d_2$.

[33] This is just one among many (Lipschitz equivalent) notions of product distance, another one being $d(x, y) := [d_1^2(x_1, y_1) + d_2^2(x_2, y_2)]^{1/2}$. Our choice is particularly convenient in certain proofs.

7.3. REMARK. - We will frequently use the following elementary facts:

(i) If d is a semi-distance on X and $t > 0$, then $L(f, t\,d) = t^{-1}L(f, d)$.

(ii) If d_1, d_2 are semi-distances on X and $d_1 \le d_2$, then $L(f, d_1) \ge L(f, d_2)$.

(iii) If d_1, d_2 are semi-distances on X and $L(f, d_1) = L(f, d_2) = L$, then $L(f, d_1 \wedge d_2) = L$.

7.4. EXAMPLE. - Let X be an interval in $\mathbb{R}$. Given a positive function ρ in $L^1(X)$, we set

$$d_\rho(x, y) := \int_{[x,y]} \rho\, d\mathscr{L}^1 \quad \text{for every } x, y \in X.$$

The function d_ρ is a semi-distance on X, and it is a distance if the essential support of ρ is X (i.e., there is no proper interval in X where ρ is a.e. null); in the latter case d_ρ is topologically equivalent to the Euclidean distance.

If ρ_1 and ρ_2 are positive functions in $L^1(X)$, then $d_{\rho_1} \wedge d_{\rho_2} = d_\rho$ where ρ is the pointwise infimum of ρ_1 and ρ_2.[34] In particular, if we choose ρ_1 and ρ_2 so that the essential support of both functions is X and $\rho = 0$ a.e. in X,[35] then d_{ρ_1} and d_{ρ_2} are topologically equivalent while the involution $d_{\rho_1} \wedge d_{\rho_2}$ is null.

7.5. DEFINITION OF $[d; d']$. - Given d and d' semi-distances on the set X, we define

$$[d; d'] := \sup_{t>0} (d \wedge t\, d')\,. \tag{7.1}$$

Being a pointwise supremum of a family of semi-distances, $[d; d']$ is a semi-distance on X, and clearly $[d; d'] \le d$. Note that the operator $[d; d']$ is not symmetric in its arguments.

From now on, d and d' are distances on a given set X.

7.6. LEMMA. - *Assume that (X, d) is compact, and let f be a function in* $\mathrm{Lip}(X, d)$. *If $L(f, [d; d']) < +\infty$, then there exists a sequence (f_n) of functions in* $\mathrm{Lip}(X, d) \cap \mathrm{Lip}(X, d')$ *that converge uniformly to f and satisfy*

$$L(f_n, d) \le L(f_n, [d; d']) \le L(f, [d; d']) \quad \textit{for every } n. \tag{7.2}$$

Conversely, if there exists a sequence (f_n) in $\mathrm{Lip}(X, d) \cap \mathrm{Lip}(X, d')$ *that converges pointwise to f, then*

$$L(f, [d; d']) \le \liminf_{n \to +\infty} L(f_n, d)\,. \tag{7.3}$$

[34] That is, $\rho(x) := \min\{\rho_1(x), \rho_2(x)\}$ for every $x \in X$. The proof of this statement requires some care, and since it is not essential to the rest of the paper, it has been omitted.

[35] For example, let ρ_1, ρ_2 be the characteristic functions of the sets E and $X \setminus E$, where E satisfies $0 < \mathscr{L}^1(E \cap I) < \mathscr{L}^1(I)$ for every interval I contained in X (note that there exists such a set).

PROOF. - Let n be a positive integer. Since X is compact w.r.t. d, it is also totally bounded, and therefore we can find a finite set E_n in X such that

$$\text{for every } x \in X \text{ there exists } x' \in E_n \text{ such that } d(x, x') \le 1/n. \tag{7.4}$$

Denote by f_n the restriction of $(1 - 1/n) \cdot f$ to the set E_n. Since E_n is finite, we can find t_n such that $d \wedge t_n d' \ge (1 - 1/n) \cdot [d; d']$ for every couple of points in E_n, and then

$$L(f_n, d \wedge t_n d') \le L(f, [d; d']) . \tag{7.5}$$

Now we use McShane's lemma to extend f_n to the rest of X without increasing the Lipschitz constant w.r.t. $d \wedge t_n d'$. Thus (7.5) holds for the extended function too. Since $d \wedge t_n d' \le [d; d'] \le d$ and $d \wedge t_n d' \le t_n d'$, inequality (7.5) implies

$$L(f_n, d) \le L(f_n, [d; d']) \le L(f_n, d \wedge t_n d') \le L(f, [d; d']) \tag{7.6}$$

and

$$\frac{1}{t_n} L(f_n, d') = L(f_n, t_n d') \le L(f_n, d \wedge t_n d') \le L(f, [d; d']) .$$

We show next that the functions f_n converge uniformly to f. Fix $x \in X$; by (7.4) there exists $x' \in E_n$ such that $d(x, x') \le 1/n$. Then $f_n(x') = (1 - 1/n) \cdot f(x')$ by construction, and taking into account (7.6) we obtain

$$\begin{aligned} |f_n(x) - f(x)| &\le |f_n(x) - f_n(x')| + |f_n(x') - f(x')| + |f(x') - f(x)| \\ &\le \frac{1}{n}\big[L(f_n, d) + \|f\| + L(f, d)\big] \le \frac{1}{n}\big[\|f\| + 2L(f, [d; d'])\big] . \end{aligned}$$

Hence $\|f_n - f\| = O(1/n)$.

Let us prove the second part of Lemma 7.6. For every n we set $L_n := L(f_n, d)$, $L'_n := L(f_n, d')$, and $t_n := L'_n / L_n$.[36] Then $L(f_n, t_n d') = L(f_n, d) = L_n$, and by Remark 7.3(iii),

$$L(f_n, d \wedge t_n d') = L_n .$$

Moreover $d \wedge t_n d' \le [d; d']$ implies $L(f_n, [d; d']) \le L(f_n, d \wedge t_n d')$, and therefore

$$\begin{aligned} L(f, [d; d']) \le \liminf_{n \to +\infty} L(f_n, [d; d']) &\le \liminf_{n \to +\infty} L(f_n, d \wedge t_n d') \\ &= \liminf_{n \to +\infty} L_n = \liminf_{n \to +\infty} L(f_n, d) . \end{aligned}$$ □

7.7. LEMMA. - *Assume that (X, d) is compact. $L(f, [d; d'])$ is finite for every f in* $\mathrm{Lip}(X, d)$ *if and only if there exists a constant $s > 0$ such that $[d; d'] \ge s\, d$.*

PROOF. - The "if" part is straightforward.

[36] We assume that f is not constant, so that $L_n \ne 0$ for n sufficiently large.

To prove the "only if" part, we fix a point $x_0 \in X$ and denote by $\mathrm{Lip}_0(X,d)$ the linear space of all $f \in \mathrm{Lip}(x,d)$ such that $f(x_0) = 0$. Then $L(f,d)$ is a Banach norm on $\mathrm{Lip}_0(X,d)$. The Banach space $\mathrm{Lip}_0(X,[d;d'])$ is defined in the same way.

The inequality $[d;d'] \leq d$ implies $L(f,[d;d']) \geq L(f,d)$, which means that the identity operator $I : \mathrm{Lip}(X,[d;d']) \to \mathrm{Lip}(X,d)$ is well-defined and bounded. Since I is surjective by assumption, the Open Mapping Theorem implies that the inverse of I is also bounded. Thus there exists a constant $s > 0$ such that $L(f,d) \geq s\, L(f,[d;d'])$ for every $f \in \mathrm{Lip}(X,d)$.

Finally, for every $y \in X$ we set $f_y(x) := d(x,y) - d(x_0,y)$. Then f_y belongs to $\mathrm{Lip}_0(x,d)$ and $L(f_y,d) = 1$. Hence $L(f_y,[d;d']) \leq 1/s$, and therefore, for every $x \in X$ there holds

$$d(x,y) = |f_y(x) - f_y(y)| \leq \frac{1}{s}[d;d'](x,y)\,. \qquad \square$$

The following result is a straightforward corollary of Lemmas 7.6 and 7.7.

7.8. Proposition. - *Assume that (X,d) is compact. If d and d' satisfy the condition*

$$[d;d'] \geq s\, d \quad \textit{for some } s > 0, \tag{7.7}$$

then for every function f in $\mathrm{Lip}(X,d)$ there exists a sequence of functions (f_n) in $\mathrm{Lip}(X,d) \cap \mathrm{Lip}(X,d')$ that converge uniformly to f and satisfy $L(f_n,d) \leq \frac{1}{s}L(f,d)$ for every n.

Conversely, if (7.7) does not hold, then there exists f in $\mathrm{Lip}(X,d)$ which cannot be approximated (not even pointwise) by any sequence of functions (f_n) in $\mathrm{Lip}(X,d) \cap \mathrm{Lip}(X,d')$ with $L(f_n,d)$ uniformly bounded.

7.9. Remarks. - (i) Condition (7.7) is obviously satisfied if d and d' are Lipschitz equivalent, but not if they are just topologically equivalent. Take indeed X, d_{ρ_1} and d_{ρ_2} as in Example 7.4, choosing the functions ρ_1 and ρ_2 so that their essential supports agree with X and their pointwise infimum is a.e. null: then d_{ρ_1} and d_{ρ_2} are topologically equivalent but $d_{\rho_1} \wedge t\, d_{\rho_2} = 0$ for every $t > 0$, and therefore $[d_{\rho_1}; d_{\rho_2}] = 0$.

(ii) Condition (7.7) is weaker than requiring $d' \geq s\, d$ for some $s > 0$. Take indeed X, d_{ρ_1} and d_{ρ_2} as in Example 7.4 with $\rho_1 = 1$ and ρ_2 so that $\rho_2 > 0$ a.e., and the essential infimum of ρ_2 is 0: then $[d_{\rho_1}; d_{\rho_2}] = d_{\rho_1}$, but there exists no $s > 0$ such that $d_{\rho_2} \geq s\, d_{\rho_1}$.

(iii) If d and d' satisfy condition (7.7), then the same is true if we replace d or d' by Lipschitz equivalent distances.

(iv) If d_1 and d_1' are distances on X_1 and d_2 is a distance on X_2, then it is easy to check that the product distances $d_1 \times d_2$ and $d_1' \times d_2$ (cf. §7.1) satisfy $[d_1 \times d_2; d_1' \times d_2] = [d_1; d_1'] \times d_2$. In particular, if d_1 and d_1' satisfy condition (7.7), then the same is true for $d_1 \times d_2$ and $d_1' \times d_2$.

(v) If (X, d') is connected by curves with finite length and d is the geodesic distance associated to d',[37] then $[d; d'] = d \wedge d' = d$.

Finally, we show that condition (7.7) is satisfied in a particular situation that occurs in the proof of Lemma 2.14. Let $X := [a, b]$ be an interval endowed with the Euclidean distance d. Given a continuous *injective* map $\gamma : X \to \mathbb{R}^d$, let d' be the pull-back according to γ of the Euclidean distance on $\mathbb{R}^d$, that is,

$$d'(x, y) := |\gamma(x) - \gamma(y)| \quad \text{for all } x, y \in X.$$

7.10. Proposition. - *Let γ be as above. If γ belongs to $W^{1,1}$ and $|\dot{\gamma}| > 0$ a.e., then $[d; d'] = d$. Thus condition (7.7) is satisfied with $s = 1$.*

7.11. Remark. - The statement of Proposition 7.10 can be easily extended to the case when X is an interval with identified end points $[a, b]^*$ (cf. §2.2) and d is the distance defined in (2.2).

In order to prove Proposition 7.10 we need some additional notation and lemmas. Let $I := [x, y]$ be an interval. A *partition* of I is a finite family $\mathscr{P}$ of closed intervals $J \subset I$ which cover I and have pairwise disjoint interiors. We write $\sigma(\mathscr{P})$ for the maximum of the diameter of all $J \in \mathscr{P}$ and denote by $V(\mathscr{P})$ the linear space of all ϕ in $L^\infty(I; \mathbb{R}^d)$ that are a.e. constant on every $J \in \mathscr{P}$.

7.12. Lemma. - *Let be given a sequence of partitions $\mathscr{P}_n$ of I such that $\sigma(\mathscr{P}_n)$ tends to 0 as $n \to +\infty$. Then for every map ϕ in $L^\infty(I; \mathbb{R}^d)$ there exists a sequence of maps $\phi_n \in V(\mathscr{P}_n)$ which converge a.e. to ϕ and satisfy $\|\phi_n\|_\infty \leq \|\phi\|_\infty$ for all n.*

Proof. - The existence of the required approximation can be easily deduced from the following statements and a suitable diagonal argument: a) for every ϕ in $L^\infty(I; \mathbb{R}^d)$ there exist continuous maps $\phi_n : X \to \mathbb{R}^d$ with $\|\phi_n\|_\infty \leq \|\phi\|_\infty$ that converge a.e. to ϕ (use Lusin theorem); b) every continuous map from X to $\mathbb{R}^d$ can be approximated uniformly by maps $\phi_n \in V(\mathscr{P}_n)$. $\square$

7.13. Lemma. - *Let $\gamma : I \to \mathbb{R}^d$ be a continuous map of class $W^{1,1}$ such that $|\dot{\gamma}| > 0$ a.e. For every positive integer n, let be given a finite family $\mathscr{F}_n$ of intervals $I_{n,i} = [x_{n,i}, y_{n,i}]$ such that*

$$\lim_{n \to +\infty} \max_i (y_{n,i} - x_{n,i}) = 0\,, \tag{7.8}$$

$$\lim_{n \to +\infty} \sum_i |\gamma(y_{n,i}) - \gamma(x_{n,i})| = 0\,. \tag{7.9}$$

For every n let A_n be the union of $I_{n,i}$ over all i. Then $\lim_{n \to +\infty} \mathscr{L}^1(A_n) = 0$.

[37] That is, $d(x, y)$ is the infimum of the length among all curves that contain x and y.

7.14. REMARK. - Assumption (7.9) yields that $\max_i |\gamma(y_{n,i}) - \gamma(x_{n,i})|$ tends to 0 as $n \to +\infty$. If γ is injective this implies assumption (7.8).

PROOF. - *Step 1.* We first prove the claim under the assumption that for every n the intervals in $\mathscr{F}_n$ have pairwise disjoint interiors.

Set $L := \limsup \mathscr{L}^1(A_n)$, and let g_n be the characteristic function of A_n. Passing to a subsequence in n we can assume that $\mathscr{L}^1(A_n)$ converge to L and the functions g_n converge to some g in the weak* topology of $L^\infty(I)$. Hence $0 \leq g \leq 1$ a.e., and

$$\int_I g \, d\mathscr{L}^1 = \lim_{n\to+\infty} \int_I g_n \, d\mathscr{L}^1 = L \,. \tag{7.10}$$

Since the intervals in $\mathscr{F}_n$ have pairwise disjoint interiors and satisfy (7.8), we can construct partitions $\mathscr{P}_n$ of I which contain $\mathscr{F}_n$ and satisfy $\sigma(\mathscr{P}_n) \to 0$.

Set $\phi := \dot\gamma/|\dot\gamma|$. By Lemma 7.12 there exists a sequence of maps ϕ_n in $V(\mathscr{P}_n)$ that converge a.e. to ϕ and satisfy $\|\phi_n\|_\infty \leq \|\phi\|_\infty = 1$. Since ϕ_n is a.e. constant on each $I_{n,i}$ there holds

$$\begin{aligned} \sum_i |\gamma(y_{n,i}) - \gamma(x_{n,i})| &= \sum_i \left| \int_{I_{n,i}} \dot\gamma \, d\mathscr{L}^1 \right| \\ &\geq \sum_i \int_{I_{n,i}} \langle \phi_n ; \dot\gamma \rangle \, d\mathscr{L}^1 = \int_I \langle g_n \phi_n ; \dot\gamma \rangle \, d\mathscr{L}^1 \,. \end{aligned} \tag{7.11}$$

The functions ϕ_n and g_n are uniformly bounded, ϕ_n converge a.e. to ϕ, and g_n converge weakly* to g. Therefore $g_n\phi_n$ converge weakly* to $g\phi$ in $L^\infty(I;\mathbb{R}^d)$ and

$$\lim_{n\to+\infty} \int_I \langle g_n\phi_n ; \dot\gamma \rangle \, d\mathscr{L}^1 = \int_I \langle g\phi ; \dot\gamma \rangle \, d\mathscr{L}^1 = \int_I g \, |\dot\gamma| \, d\mathscr{L}^1 \,. \tag{7.12}$$

Putting together (7.9), (7.11), and (7.12) we finally obtain $\int_I g \, |\dot\gamma| \, d\mathscr{L}^1 = 0$. Since $g \geq 0$ and $|\dot\gamma| > 0$ a.e., we infer that $g = 0$ a.e., and then (7.10) implies L=0.

Step 2. We now remove the assumption that the intervals in $\mathscr{F}_n$ have pairwise disjoint interiors. Given an interval $J = [x-d, x+d]$, we denote by $\hat J$ the concentric interval $[x-5d, x+5d]$. By a standard covering argument (cf. [19], Theorem 3.3), we can extract from each $\mathscr{F}_n$ a subfamily $\mathscr{F}'_n$ of pairwise disjoint intervals J such that the corresponding $\hat J$ cover A_n. Hence, $\mathscr{L}^1(A_n) \leq 5 \, \mathscr{L}^1(A'_n)$ where A'_n is the union of the intervals in $\mathscr{F}'_n$, and we know from Step 1 that $\mathscr{L}^1(A'_n) \to 0$. □

PROOF OF PROPOSITION 7.10. - We know that $[d;d'] \leq d$, and assume by contradiction that there exist $x < y$ in X such that $[d;d'](x,y) < d(x,y)$.

Choose L such that $[d;d'](x,y) < L < d(x,y)$. Since $[d;d']$ is the supremum of $d \wedge td'$, recalling the definitions of d, d', and $d \wedge td'$, we obtain that for every

$t > 0$ there exist finitely many points $\{z_{t,0}, \dots, z_{t,n_t}\}$ in X such that $z_{t,0} = x$, $z_{t,n_t} = y$, and

$$S_t := \sum_{i=1}^{n_t} \min\left\{|z_{t,i-1} - z_{t,i}|\,,\, t|\gamma(z_{t,i-1}) - \gamma(z_{t,i})|\right\} \le L\,. \tag{7.13}$$

For every t, let $\mathscr{I}_t$ be the set of all indices i such that the minimum in (7.13) agrees with $t|\gamma(z_{t,i-1}) - \gamma(z_{t,i})|$. Then (7.13) implies that

$$\lim_{t\to+\infty} \sum_{i\in\mathscr{I}_t} |\gamma(z_{t,i-1}) - \gamma(z_{t,i})| = 0\,.$$

Hence the families of all intervals $[z_{t,i-1}, z_{t,i}]$ with $i \in \mathscr{I}_t$ satisfy the assumptions of Lemma 7.13 (recall Remark 7.14), and therefore $\mathscr{L}^1(A_t) \to 0$ as $t \to +\infty$, where A_t is the union of such intervals. Since

$$S_t \ge \sum_{i\notin\mathscr{I}_t} |z_{t,i-1} - z_{t,i}| \ge \mathscr{L}^1(X \setminus A_t) = y - x - \mathscr{L}^1(A_t)\,,$$

we deduce that

$$\limsup_{t\to+\infty} S_t \ge y - x\,,$$

and this contradicts (7.13) because $y - x = d(x, y) > L$ by the choice of L. □

We conclude this section with an extension of Proposition 7.10 used in [1].

7.15. COROLLARY. - *Let $X := [a_1, b_1] \times [a_2, b_2]$ be a closed rectangle endowed with the Euclidean distance d, and let d' be the pull-back of the Euclidean distance on $\mathbb{R} \times \mathbb{R}^d$ according to the map $(x_1, x_2) \mapsto (x_1, \gamma(x_2))$, where γ is taken as in Proposition 7.10. Then d and d' satisfy condition* (7.7).

PROOF. - Let d_1 and d_2 be the Euclidean distances on $[a_1, b_1]$ and $[a_2, b_2]$, and let d_2' be the pull-back of the Euclidean distance on $\mathbb{R}^d$ according to γ. By Proposition 7.10, d_2 and d_2' satisfy condition (7.7). Then, as pointed out in Remark 7.9(iv), also the product distances $d_1 \times d_2$ and $d_1 \times d_2'$ satisfy (7.7). To conclude, it suffices to observe that d and d' are Lipschitz equivalent to $d_1 \times d_2$ and $d_1 \times d_2'$ and recall Remark 7.9(iii). □

7.16. REMARK. - Like Proposition 7.10, also Corollary 7.15 remains valid when the second factor $[a_2, b_2]$ in the product X is replaced by an interval with identified end points $[a_2, b_2]^*$, provided that X is endowed with a distance d which is Lipschitz equivalent to the product distance.

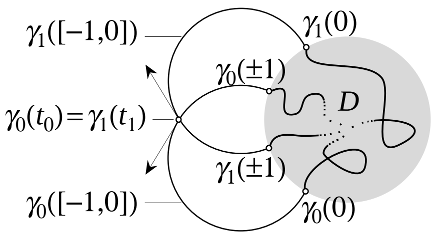

Figure 8. Proof of Lemma 8.1: extension of paths γ_1 and γ_2.

8. Appendix: A proof of the triod lemma

In this appendix we give a self-contained proof of Lemma 2.15.

8.1. Lemma. - *Let D be a closed disc in $\mathbb{R}^2$, and let C_1, C_2 be curves parametrized by $\gamma_1, \gamma_2 : [0,1] \to D$ with end points $\gamma_i(0), \gamma_i(1)$ in ∂D for $i = 1, 2$. If C_1 and C_2 are disjoint, then $\gamma_2(0)$ and $\gamma_2(1)$ belong to the same connected component of $\partial D \setminus \{\gamma_1(0), \gamma_1(1)\}$.*

Proof. - Assume by contradiction that $\gamma_2(0)$ belongs to one of the two arcs of the circle ∂D with end points $\gamma_1(0)$ and $\gamma_1(1)$ while $\gamma_2(1)$ belongs to the other. Then we can extend both paths γ_1, γ_2 to $[-1,1]^*$ so that the extensions are smooth on $[-1,0]$, map the interval $(-1,0)$ in the complement of D, and there exists only one couple (t_1, t_2) such that $\gamma_1(t_1) = \gamma_2(t_2)$, and in that case the vectors $\dot\gamma_1(t_1)$ and $\dot\gamma_2(t_2)$ are linearly independent (see Figure 8).

Then $\Gamma(s_1, s_2) := \gamma_1(s_1) - \gamma_2(s_2)$ is a map from the torus $[-1,1]^* \times [-1,1]^*$ to $\mathbb{R}^2$, and the assumptions on γ_1 and γ_2 imply that $\Gamma^{-1}(0)$ contains only one point, and this point is regular. Hence the Brower degree of Γ must be ± 1. On the other hand Γ is not surjective on $\mathbb{R}^2$ because its domain is compact, and therefore its Brower degree must be 0. □

8.2. Corollary. - *Let D be a closed disc in $\mathbb{R}^2$, and let E_1 and E_2 be close subsets of D which are connected by curves. If E_1 and E_2 are disjoint, then $E_1 \cap \partial D$ is contained in one connected component of $\partial D \setminus E_2$.*

Proof. - Assume by contradiction that there exist $x_1, y_1 \in E_1 \cap \partial D$ which do not belong to the same connected component of $\partial D \setminus E_2$. Then the two arcs of ∂D with end points x_1 and y_1 must both contain points of E_2, say x_2 and y_2. Thus x_2 and y_2 belongs to different connected components of $\partial D \setminus \{x_1, y_1\}$. But for $i = 1, 2$ there exist curves C_i contained in E_i and connecting x_i and y_i, and this contradicts Lemma 8.1. □

Proof of Lemma 2.15. - Let D be an open disc in $\mathbb{R}^2$ with center x and radius ρ. For every triod $Y = C_1 \cup C_2 \cup C_3$ with center y we set

$$r := \min_{i=1,2,3} \Big\{ \max_{z \in C_i} |z - y| \Big\} .$$

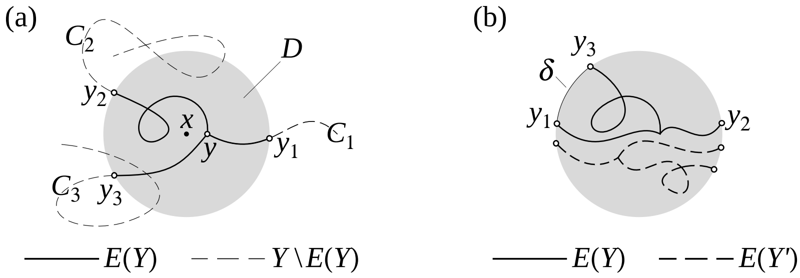

Figure 9. Proof of Lemma 2.15: (a) truncation of the triod Y; (b) the distance between $F(Y)$ and $F(Y')$ is larger than δ.

If $y \in D$ and $r \geq |x-y| + \rho$, then for every i there exists $z \in C_i$ which does not belong to D, and therefore C_i must intersect the circle ∂D. Then we denote by t_i the smallest t such that $\gamma_i(t) \in \partial D$ (we follow the notation of §2.3), and by $E(Y)$ the 'truncated' triod

$$E(Y) := \bigcup_{i=1}^{3} \gamma_i([a_i, t_i])$$

(see Figure 9(a)). Moreover we let $F(Y)$ be the set of end points of $E(Y)$, i.e.,

$$F(Y) := \{\gamma_i(t_i) : i = 1, 2, 3\} = E(Y) \cap \partial D\,.$$

Finally, let $\mathscr{F}_D$ denote the family of all $Y \in \mathscr{F}$ such that $y \in D$ and $r \geq |x-y| + \rho$, and $\mathscr{G}_D$ the family of all $F(Y)$ with $Y \in \mathscr{F}_D$.

Step 1. The family $\mathscr{F}$ is the union of $\mathscr{F}_D$ over all discs D whose centers have rational coordinates, and whose radii are rational. Hence, to prove that $\mathscr{F}$ is countable it suffices to show that each $\mathscr{F}_D$ is countable.

Step 2. Since the elements of $\mathscr{F}_D$ are pairwise disjoint by assumption and $F(Y) \subset Y$ for every Y, the map $Y \mapsto F(Y)$ is injective on $\mathscr{F}_D$. Therefore to prove that $\mathscr{F}_D$ is countable it suffices to show that $\mathscr{G}_D$ is countable.

Step 3. On the circle ∂D we consider the geodesic distance d,[38] and endow the set $\mathscr{G}_D$ with the Hausdorff distance d_H associated to d (see (2.1)). Since the metric space $\mathscr{G}_D$ is separable, to prove that it is countable it suffices to show that it contains only isolated points.

Step 4. Given $Y \in \mathscr{F}_D$, let δ be the length of the shortest arc in ∂D with end points in $F(Y) = \{y_1, y_2, y_3\}$ (see Figure 9(b)). We claim that $d_H(F(Y), F(Y')) \geq \delta$ for every $Y' \in \mathscr{F}_D$ with $Y' \neq Y$. Indeed the triods $E(Y)$ and $E(Y')$ are disjoint subsets of the closure of D, and Corollary 8.2 implies that $E(Y') \cap \partial D = F(Y')$ is contained in one connected component of $\partial D \setminus E(Y) = \partial D \setminus F(Y)$, for example the arc with end points y_1, y_2 that does not contains y_3. Then the distance of y_3 from $F(Y')$ is at least δ, which implies the claim. □

[38] That is, $d(x, y)$ is the length of the shortest among the two arcs of ∂D with end points x, y.

References

1. G. Alberti, S. Bianchini, G. Crippa: A uniqueness result for the continuity equation in two dimensions. In preparation.
2. S.M. Bates: Toward a precise smoothness hypothesis in Sard's theorem. *Proc. Amer. Math. Soc.* **117** (1993), no. 1, 279–283.
3. B. Bojarski, P. Hajłasz, P. Strzelecki: Sard's theorem for mappings in Hölder and Sobolev spaces. *Manuscripta Math.* **118** (2005), no. 3, 383–397.
4. L. De Pascale: The Morse-Sard theorem in Sobolev spaces. *Indiana Univ. Math. J.* **50** (2001), no. 3, 1371–1386.
5. R.J. DiPerna, P.-L. Lions: Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** (1989), no. 3, 511–547.
6. A.Ya. Dubovickiĭ: On the structure of level sets of differentiable mappings of an n-dimensional cube into a k-dimensional cube (Russian). *Izv. Akad. Nauk SSSR. Ser. Mat.* **21** (1957), 371–408.
7. R. Engelking: *General topology. Revised and completed edition.* Sigma Series in Pure Mathematics, 6. Heldermann Verlag, Berlin 1989.
8. K.J. Falconer: *The geometry of fractal sets.* Cambridge Tracts in Mathematics, 85. Cambridge University Press, Cambridge 1985.
9. H. Federer: *Geometric measure theory.* Grundlehren der mathematischen Wissenschaften, 153. Springer, Berlin-New York 1969. Reprinted in the series Classics in Mathematics. Springer, Berlin-Heidelberg 1996.
10. A. Figalli: A simple proof of the Morse-Sard theorem in Sobolev spaces. *Proc. Amer. Math. Soc.* **136** (2008), no. 10, 3675–3681.
11. D.H. Fremlin: Spaces of finite length. *Proc. London Math. Soc. (3)* **64** (1992), no. 3, 449–486.
12. E.L. Grinberg: On the smoothness hypothesis in Sard's theorem. *Amer. Math. Monthly* **92** (1985), no. 10, 733–734.
13. M.W. Hirsch: *Differential topology. Corrected reprint of the 1976 original.* Graduate Texts in Mathematics, 33. Springer-Verlag, New York 1976.
14. S.V. Konyagin: On the level sets of Lipschitz functions. *Tatra Mt. Math. Publ.* **2** (1993), 51–59.
15. S.G. Krantz, H.R. Parks: *Geometric integration theory.* Cornerstones. Birkhäuser, Boston 2008.
16. R.L. Moore: Concerning triods in the plane and the junction points of plane continua. *Proc. Nat. Acad. Sci. U.S.A.* **14** (1928), no. 1, 85-88.
17. C.R. Pittman: An elementary proof of the triod theorem. *Proc. Amer. Math. Soc.* **25** (1970), no. 4, 919.
18. A. Sard: The measure of the critical values of differentiable maps. *Bull. Am. Math. Soc.* **48** (1942), 883–890.
19. L. Simon: *Lectures on geometric measure theory.* Proceedings of the Centre for Mathematical Analysis, 3. Australian National University, Centre for Mathematical Analysis, Canberra 1983.
20. H. Whitney: A function not constant on a connected set of critical points. *Duke Math. J.* **1** (1935), 514–517.

G.A.: Dipartimento di Matematica, Università di Pisa, largo Pontecorvo 5, 56127 Pisa, Italy.
e-mail: `galberti1@dm.unipi.it`

S.B.: S.I.S.S.A., via Beirut 4, 34014 Trieste, Italy.
e-mail: `bianchin@sissa.it`

G.C.: Dipartimento di Matematica, Università degli Studi di Parma, viale G.P. Usberti 53/A (Campus), 43124 Parma, Italy.
e-mail: `gianluca.crippa@unipr.it`