A NOTE ON THE DIFFERENTIABILITY OF LIPSCHITZ FUNCTIONS AND THE CHAIN RULE IN SOBOLEV SPACES

M. MORINI

ABSTRACT. We prove a new differentiability criterion for Lipschitz functions $f : \mathbb{R}^d \to \mathbb{R}$ and we discuss its implications on the validity of the chain rule for the weak derivative of $f \circ u$, when u is a vector-valued Sobolev function.

Keywords: differentiability of Lipschitz functions, chain rule, Sobolev functions

1. Introduction

In this note we provide a new differentiability criterion for Lipschitz functions, which allows to settle a question related to the validity of the classical chain rule for the composite function $f \circ u$, when u ia a vector-valued Sobolev map and f is Lispchitz continuous. Before stating the results we recall that a Borel set $E \subset \mathbb{R}^d$ is said *(countably)* $\mathcal{H}^k$*-rectifiable*, $k \in \{1, \ldots, d-1\}$, if there exists a sequence $\{M_n\}$ of k-dimensional C^1 manifolds such that

$$\mathcal{H}^k\left(E \setminus \bigcup_{n=1}^{\infty} M_n\right) = 0.$$

Here $\mathcal{H}^k$ denotes the k-dimensional Hausdorff measure. We also recall that $\mathcal{H}^k$-a.e. $w \in E$ admits an *approximate (k-dimensional) tangent space* $\mathrm{Tan}^k(E, w)$ (see [2] for more details). For all such w's we let $\pi_w^E : \mathbb{R}^d \to \mathbb{R}^d$ denote the orthogonal projection onto $\mathrm{Tan}^k(E, w)$.

We will prove the following differentiability criterion for Lipschitz functions.

Theorem 1.1. *Let $f : \mathbb{R}^d \to \mathbb{R}$, $d \geq 3$, be Lipschitz continuous and let $E \subset \mathbb{R}^d$ be an $\mathcal{H}^k$-rectifiable set, with $1 \leq k \leq d-2$. Assume that there exist a bounded function $g : E \to \mathbb{R}^d$ and a countable dense family of orthonormal bases $\mathcal{S}$ such that for every $\{e_1, \ldots, e_d\} \in \mathcal{S}$ the gradient $\nabla^{(e_1 \ldots e_d)} f := \sum_{i=1}^d \frac{\partial f}{\partial e_i} e_i$ exists $\mathcal{H}^k$-a.e. in E and satisfies*

$$\pi_w^E\left(\nabla^{(e_1 \ldots e_d)} f(w)\right) = g(w) \quad \text{for } \mathcal{H}^k\text{-a.e. } w \in E\,. \tag{1.1}$$

Then f is differentiable $\mathcal{H}^k$-a.e. in E.

It is well-known that if $\nabla^{(e_1 \ldots e_d)} f(w)$ is the same vector for all orthonormal bases $\{e_1, \ldots, e_d\}$ in countable dense collection $\mathcal{S}$, then f is differentiable at w (see Proposition 2.1 below). Condition (1.1), instead, states that only the tangential part of $\nabla^{(e_1 \ldots e_d)} f(w)$ is invariant as $\{e_1, \ldots, e_d\}$ ranges over $\mathcal{S}$. Note that such an invariance condition is not enough to guarantee the differentiability at a given point w: Indeed, it is possible to construct a Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$ not differentiable at the origin and such that $\nabla^{(e_1 \ldots e_d)} f(0)$ is parallel to a given vector for *all* othonormal bases $\{e_1, \ldots, e_d\}$ (see Subsection 2.1). In view of this example one has to fully exploit the rectifiable structure of E to show that, in fact, condition (1.1) yields differentiability $\mathcal{H}^k$-a.e. along E. Clearly, a posteriori, the function g in (1.1) will coincide $\mathcal{H}^k$-a.e. with the tangential gradient of f along E .

The restriction $k \leq d-2$ is motivated by the fact that if E is $\mathcal{H}^{d-1}$-rectifiable, then it suffices to assume partial differentiabilty along nontangential directions in order to guarantee full differentiability $\mathcal{H}^{d-1}$-a.e. on E. More precisely, the following holds (see [7, Theorem 3.1 and Remark 3.2]): *If $E \subset \mathbb{R}^d$ is $\mathcal{H}^{d-1}$-rectifiable and for some $\sigma \in S^{d-1}$ the partial derivative $\frac{\partial f}{\partial \sigma}$ exists $\mathcal{H}^{d-1}$-a.e. on E, then f is differentiable $\mathcal{H}^{d-1}$-a.e. on the set $\{w \in E : \sigma \notin$*

$\mathrm{Tan}^{d-1}(E,w)\}$. The situation is completely different when $k < d-1$. Indeed, as shown in [7, Theorem 4.10 and Remark 4.11(ii)], for any given $\mathcal{H}^k$-rectifiable set E, with $k < d-1$, it is possible to construct a Lipschitz function that admits *all* directional derivatives at every point, and yet the singular set

$$\Sigma^f := \left\{ w \in \mathbb{R}^d : f \text{ is not differentiable at } w \right\} \tag{1.2}$$

coincides, up to an $\mathcal{H}^k$-negligible set, with E. Moreover, a careful inspection of the proof shows that the function constructed in the quoted theorem satisfies (1.1) with respect to an infinite nondense collection of orthonormal frames, which shows that the density assumption in Theorem 1.1 cannot be removed.

We now illustrate the link between the differentiability criterion provided in Theorem 1.1 and the validity of the chain rule for $f \circ u$ with $u \in W^{1,1}_{\mathrm{loc}}(\Omega;\mathbb{R}^d)$. Since the problem is local, we assume, without loss of generality, that $\Omega = \mathbb{R}^N$. Let $d > 1$ and let $\{e_1,\ldots,e_d\}$ be any fixed basis of $\mathbb{R}^d$, not necessarily orthonormal. We say that the *classical chain rule* holds for the composite function $f \circ u$ if[1]

$$\frac{\partial}{\partial x_j}(f \circ u) = \sum_{i=1}^{d} \frac{\partial f}{\partial e_i}(u)\,\frac{\partial u_i}{\partial x_j} \quad \mathcal{L}^N\text{-a.e. in } \mathbb{R}^N, \tag{1.3}$$

where $\frac{\partial f}{\partial e_i}(u)\,\frac{\partial u_i}{\partial x_j}(x)$ is interpreted to be zero whenever $\frac{\partial u_i}{\partial x_j} = 0$, irrespective of whether $\frac{\partial f}{\partial e_i}(u)$ is defined. The main obstruction to the validity of the chain rule lies in the fact that the right-hand side of (1.3) may be nowhere defined, as simple examples show.

In [1] Ambrosio and Dal Maso proved a weaker but very general form of the chain rule. They showed that for every $u \in W^{1,1}_{\mathrm{loc}}\left(\mathbb{R}^N;\mathbb{R}^d\right)$ the restriction of f to the affine space

$$T^u_x := \left\{ w \in \mathbb{R}^d : w = u(x) + \nabla u(x)\, z \text{ for some } z \in \mathbb{R}^N \right\}$$

is differentiable at $u(x)$ and

$$\nabla(f \circ u)(x) = \nabla_u(f|_{T^u_x})(u(x))\,\nabla u(x)\,. \tag{1.4}$$

for $\mathcal{L}^N$-a.e. $x \in \mathbb{R}^N$. However, formula (1.4) leaves us with the *open problem* of establishing under which additional conditions on the function f the right-hand side of (1.4) coincides with the right-hand side of (1.3), in other words, to *find necessary and sufficient conditions on f for the classical chain rule* (1.3) *to hold.*

A thorough investigation of the relation between the validity of (1.3) and the structure of the singular set Σ^f of f has been carried out in [7], to which we refer for a full account of the results (see also [8, 9, 10] and the more recent paper [4] for some previous important work in this direction). Here we just recall the following three facts proven in [7], which are relevant for our discussion.

(i) If $d = 2$ then (1.3) holds with respect to a fixed basis $\{e_1, e_2\}$ of $\mathbb{R}^2$ if and only if the rectifiable part of Σ^f, if nonempty, is one-dimensional and everywhere parallel to e_1 or e_2. More precisely, for every $\mathcal{H}^1$-rectifiable set $E \subset \Sigma^f$ and for $\mathcal{H}^1$ a.e. $w \in E$ either

$$\mathrm{Tan}^1(E,w) = \mathrm{span}\{e_1\} \text{ or } \mathrm{Tan}^1(E,w) = \mathrm{span}\{e_2\}\,.$$

(ii) If $d > 2$ then the validity of (1.3) gives us some information only about the $\mathcal{H}^{d-1}$-rectifiable part of Σ^f: for every $\mathcal{H}^{d-1}$-rectifiable set $E \subset \Sigma^f$ and for $\mathcal{H}^{d-1}$ a.e. $w \in E$ the approximate tangent space $\mathrm{Tan}^{d-1}(E,w)$ is a coordinate hyperplane. On the other hand, no conditions on the $\mathcal{H}^k$-structure, $1 \le k \le d-2$, of the singular set can be inferred from the validity of (1.3) with respect to a given basis. Indeed, for any orthonormal basis $\{e_1,\ldots,e_d\}$ and any $\mathcal{H}^k$-rectifiable set $E \subset \mathbb{R}^d$, with $1 \le k \le d-2$, there exists a Lipschitz function f whose singular set Σ^f coincides, up to an $\mathcal{H}^k$-negligible set, with E, and for which (1.3) holds in $W^{1,1}_{\mathrm{loc}}\left(\mathbb{R}^N;\mathbb{R}^d\right)$ with respect to $\{e_1,\ldots,e_d\}$.

[1] Here for every $u \in \mathbb{R}^d$ we write $u = u_1e_1 + \ldots + u_de_d$.

(iii) More can be said if we allow the coordinate system to vary: If Σ^f contains a nontrivial $\mathcal{H}^1$-rectifiable set, then it is possible to find a basis $\{e_1, \ldots, e_d\}$, *not necessarily orthonormal*, for which the chain rule (1.3) fails. Hence, a necessary and sufficient condition for the classical chain rule to hold with respect to *every* coordinate system is that Σ^f is *purely $\mathcal{H}^1$-unrectifiable*; i.e., $\mathcal{H}^1(\Sigma^f \cap E) = 0$ for every $\mathcal{H}^1$-rectifiable set $E \subset \mathbb{R}^d$.

The differentiability criterion proved in this paper allows to strengthen the result contained in (iii), by showing that if the rectifiable part of the singular set Σ^f is nonempty, then we may find infinitely many *orthonormal* frames for which the chain rule fails. More precisely, we can show the following theorem that improves [7, Theorem 1.5].

Theorem 1.2. *Let $\mathcal{S}$ be a countable dense set of orthonormal bases and let $f\colon \mathbb{R}^d \to \mathbb{R}$ be Lipschitz continuous, with Σ^f not purely $\mathcal{H}^1$-unrectifiable. Then there exists an orthonormal basis $\{e_1, \ldots, e_d\} \in \mathcal{S}$ such that for every $N \in \mathbb{N}$ the chain rule* (1.3) *fails for some function $u \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^N; \mathbb{R}^d)$ and the set of orthonormal frames for which* (1.3) *fails is open.*

Hence, a necessary and sufficient condition for the classical chain rule (1.3) *to hold in $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^N; \mathbb{R}^d)$ with respect to a countable dense family of orthonormal bases is that Σ^f is purely $\mathcal{H}^1$-unrectifiable.*

Let us underline that the sufficiency part of the last statement was already known (see [8, 9, 10]), while the necessity part, to the best of our knowledge, is new.

If we restrict our attention to the smaller class $\mathcal{A}_k\left(\mathbb{R}^N; \mathbb{R}^d\right)$, $k \le \min\{N, d\}$, of all functions u in $W^{1,1}_{\mathrm{loc}}\left(\mathbb{R}^N; \mathbb{R}^d\right)$ such that rank $(\nabla u(x))$ is either zero or greater than or equal to k for $\mathcal{L}^N$-a.e. $x \in \mathbb{R}^N$, then the appropriate necessary and sufficient condition is the pure $\mathcal{H}^k$-unrectifiability of the singular set Σ^f. More precisely, as a consequence of Theorem 1.1 we have the following.

Theorem 1.3. *A necessary and sufficient condition for the classical chain rule* (1.3) *to hold in $A_k(\mathbb{R}^N; \mathbb{R}^d)$, $k \le \min\{N, d\}$, with respect to a countable dense family of orthonormal bases is that Σ^f is purely $\mathcal{H}^k$-unrectifiable.*

The last theorem improves [7, Theorem 1.6] that contains the proof of the sufficiency part of the statement.

We conclude this introduction by outlining the strategy of the proof of Theorem 1.1: We start by considering the case where E is contained in a straight line of $\mathbb{R}^3$. The treatment of this case is the crucial and most difficult part of the proof and Subsection 2.3 is entirely devoted to it. Then, an induction argument on the dimension d of the ambience space allows to prove the statement when E is contained in a straight line of $\mathbb{R}^d$. If E is now any 1-rectifiable set in $\mathbb{R}^d$, we conclude by reducing to the previous case via suitable diffeomorphisms constructed by means of the Whitney Extension Theorem. Finally, a slicing argument is used to treat the general case.

2. Proof of the results

Throughout the whole section, given a unit vector $\nu \in \mathbb{R}^d$, we denote the hyperplane $\{w \in \mathbb{R}^d : w \cdot \nu = 0\}$ by Π_ν and we let $\pi_\nu : \mathbb{R}^d \to \Pi_\nu$ be the orthogonal projection onto Π_ν. We will also use the symbol S^{d-1} to indicate the unit sphere in $\mathbb{R}^d$.

The proof of Theorems 1.1 and 1.2 will be split into several subsections. The outline is the following. In the first subsection we start with the example announced in the Introduction showing that the validity of condition (1.1) at a given point w does not yield in general differentiabilty at the same point w. Hence, only by exploting the $\mathcal{H}^k$-rectifiable structure of E we can obtain $\mathcal{H}^k$-a.e. differentiability. In the second subsection we state some preliminary differentiability results for Lipschitz functions, while in the third one we collect some crucial auxiliary results concerning the case where $d = 3$ and E is contained in a straight line. In the last subsection we prove Theorems 1.1, 1.2, and 1.3 in the general case.

2.1. **An example.** In this subsection we construct a Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$ non-differentiable at the origin and such that $\nabla^{(w_1 \cdots w_d)} f(0)$ is parallel to a given vector for all othonormal bases $\{w_1, \dots, w_d\}$. To this aim fix any unit vector $\tau \in S^{d-1}$ and choose an orthonormal basis $\{e_1, \dots, e_d\}$ such that $e_d = \tau$. Denote $C := S^{d-1} \cup \{w \in R^d : w \cdot e_i \geq 0 \text{ for } i = 1, \dots, d\} \cup \{w \in R^d : w \cdot e_i \leq 0 \text{ for } i = 1, \dots, d\}$. Let $g : C \to [0, 1]$ be a non-constant even C^∞-function which vanishes on the relative boundary of C. We now define $f : \mathbb{R}^d \to \mathbb{R}$ as follows: for all $w \in \{w \in R^d : w \cdot e_i \geq 0 \text{ for } i = 1, \dots, d\} \cup \{w \in R^d : w \cdot e_i \leq 0 \text{ for } i = 1, \dots, d\}$ we set $f(w) := g(\frac{w}{|w|}) e_d \cdot w$ and $f(z) := g(\frac{w}{|w|}) e_d \cdot z$ for all $z \in \Pi_w$, where Π_w denote the hyperplane orthogonal to w. It is easy to check that f is a Lipschitz 1-homogeneous function that admits all directional derivatives at the origin. Moreover, if $\{w_1, \dots, w_d\}$ is any orthonormal basis, then one and only one vector, say w_1, belongs to C. By construction, we then have $\nabla^{(w_1, \dots, w_d)} f(0) = g(w_1) e_d$, which shows, in particular, that f is not differentiable at 0 and that the gradient is always parallel to the direction e_d.

2.2. **Preliminary differentiability results.** The first proposition states a simple differentiability criterion for Lipschitz functions whose elementary proof is given for the reader's convenience.

Proposition 2.1. *Let $f : \mathbb{R}^d \to \mathbb{R}$, $d \geq 3$, be a Lipschitz function and let $w \in \mathbb{R}^d$ be fixed. Then the following conditions are equivalent:*

(1) *f is differentiable at w;*
(2) *there exists a linear operator $L : \mathbb{R}^d \to \mathbb{R}$ such that the limit*
$$\lim_{h \to 0^+} \frac{f(w + h\nu) - f(w)}{h} = L(\nu)$$
exists for all ν in a countable dense subset of S^{d-1};
(3) *the restriction of f to every $(d-1)$-dimensional affine subspace passing through w is differentiable at w.*

Remark 2.2. *The equivalence between (1) and (2) holds also when $d = 2$.*

Proof. (1) $\Rightarrow$ (3): Obvious.

(3) $\Rightarrow$ (2): Fix an orthonormal basis $\{e_1, \dots, e_d\}$ and $\nu = \sum_{i=1}^d \lambda_i e_i \in S^{d-1}$. Set $\nu_1 := \sum_{i=1}^{d-1} \lambda_i e_i$ and $\hat{\nu}_1 = \frac{\nu_1}{|\nu_1|}$. Since f restricted to $w + \text{span}\{\hat{\nu}_1, e_d\}$ is differentiable at w, we have $\frac{\partial f}{\partial \nu}(w) = |\nu_1| \frac{\partial f}{\partial \hat{\nu}_1}(w) + \lambda_d \frac{\partial f}{\partial e_d}(w)$. In turn, the differentiability of f restricted to $w + \text{span}\{e_1, \dots, e_{d-1}\}$ yields $\frac{\partial f}{\partial \hat{\nu}_1}(w) = \frac{1}{|\nu_1|} \sum_{i=1}^{d-1} \lambda_i \frac{\partial f}{\partial e_i}(w)$. Hence, $\frac{\partial f}{\partial \nu}(w) = \nabla^{(e_1 \dots e_d)} f(w) \cdot \nu$ for every $\nu \in S^{d-1}$.

(2) $\Rightarrow$ (1): It suffices to observe that differentiability at w is equivalent to saying that for every sequence $h_n \searrow 0$ the functions $g_n(\cdot) := \frac{f(w + h_n \cdot) - f(x)}{h_n}$ converge uniformly on S^{d-1} to some linear operator. Indeed, the Lipschitz continuity of f imply that the g_n's are in turn Lipschitz continuous with the same Lipschitz constant. It follows from the assumption the $g_n \to L$ uniformly on S^{d-1}. □

Proposition 2.3. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be Lipschitz continuous and let $w \in \mathbb{R}^d$ be fixed. Assume that the restriction of f to $w + P_n$ is differentiable at w for a sequence $\{P_n\}$ of k-dimensional subspaces converging to some k-dimensional subspace P, with $k < d$. Then the restriction of f to $w + P$ is differentiable at w and its gradient is the limit of the gradients of f restricted to the approximating subspaces.*

Proof. Denoting the gradient of f restricted to $w + P_n$ by $\nabla^n f$, we may extract a subsequence (not relabelled) such that $\nabla^n f(w) \to v$ for some $v \in P$. Fix $\nu \in S^{d-1} \cap P$ and choose $\{\nu_n\} \subset S^{d-1} \cap P_n$ such that $\nu_n \to \nu$. By passing to a further subsequence, if needed, we may also assume that $\frac{\partial f}{\partial \nu_n}(w)$ converges, so that

$$\ell := \lim_{n\to\infty} \frac{\partial f}{\partial \nu_n}(w) = \lim_{n\to\infty} \nabla^n f(w)\cdot \nu_n = v\cdot \nu. \tag{2.1}$$

Denoting the Lipschitz constant of f by M, we have

$$\begin{aligned}\limsup_{t\to 0}\Big|\frac{f(w+t\nu)-f(w)}{t}-\ell\Big| &\le \limsup_{t\to 0}\Big|\frac{f(w+t\nu_n)-f(w)}{t}-\ell\Big| + M|\nu-\nu_n| \\ &= \Big|\frac{\partial f}{\partial \nu_n}(w)-\ell\Big| + M|\nu-\nu_n|.\end{aligned}$$

Let now $n\to\infty$ in the above inequality to obtain $\frac{\partial f}{\partial \nu}(w)=\ell$ and, in turn, by (2.1), $\frac{\partial f}{\partial \nu}(w) = v\cdot\nu$ for every $\nu\in S^{d-1}\cap P$. The conclusion follows from Proposition 2.1-(ii). □

We have the following immediate corollary.

Corollary 2.4. *Under the hypotheses of Theorem 1.1, condition* (1.1) *holds with respect to any orthonormal basis.*

Remark 2.5. In particular, under the hypotheses of Theorem 1.1, for $\mathcal{H}^k$-a.e. $w\in E$ the directional derivative $\frac{\partial f}{\partial \nu}(w)$ exists for all $\nu\in S^{d-1}$ and satisfies

$$\Big|\frac{\partial f}{\partial \nu_1}(w)-\frac{\partial f}{\partial \nu_2}(w)\Big| \le M|\nu_1-\nu_2|, \tag{2.2}$$

where M is the Lipschitz constant of f.

The last proposition contains a deeper differentiability result that was proven in [7, see Theorem 3.1 and Remark 3.2] and inspired by recent work by Bessis and Clarke [3].

Proposition 2.6. *Let $f:\mathbb{R}^d\to\mathbb{R}$ be Lipschitz continuous. Assume that $M\subset\mathbb{R}^d$ is $\mathcal{H}^k$-rectifiable and Π is a $(d-k)$-plane such that for $\mathcal{H}^k$-a.e. $w\in M$ the function f restricted to $w+\Pi$ is differentiable at w and*

$$\mathrm{Tan}^k(M,w)+\Pi=\mathbb{R}^d.$$

Then f is differentiable at $\mathcal{H}^k$-a.e. $w\in M$.

2.3. **Some auxiliary results.** Throughout the whole subsection we assume that $d=3$ and E is a set of positive $\mathcal{H}^1$-measure contained in a straight line; i.e.,

$$\mathcal{H}^1(E)>0 \quad\text{and}\quad E\subset \mathrm{span}\{\tau\} \text{ for some } \tau\in S^2. \tag{2.3}$$

Remark 2.7. Let $f:\mathbb{R}^3\to\mathbb{R}$ be a Lipschitz continuous function satisfying the hypotheses of Theorem 1.1 and fix any $\sigma\in S^2$ linearly independent of τ. By Remark 2.5 and by Proposition 2.6 it follows that the restriction of f to $f(0)+\mathrm{span}\{\sigma,\tau\}$ is differentiable $\mathcal{H}^1$-a.e. on E. Hence, recalling that by Corollary 2.4 condition (1.1) holds also with respect to the orthonormal basis $\{\sigma,\tau,\nu\}$, with $\nu\perp\mathrm{span}\{\sigma,\tau\}$, we deduce that g coincides $\mathcal{H}^1$-a.e on E with the tangential derivative $\frac{\partial f}{\partial \tau}$. Condition (1.1) and Corollary 2.4 then imply that

$$\frac{\partial f}{\partial \tau} = (\tau\cdot e_1)\frac{\partial f}{\partial e_1} + (\tau\cdot e_2)\frac{\partial f}{\partial e_2} + (\tau\cdot e_3)\frac{\partial f}{\partial e_3} \quad \mathcal{H}^1\text{-a.e on } E \tag{2.4}$$

for all orthonormal bases $\{e_1,e_2,e_3\}$.

Observe now that if f is differentiable along E, then for every basis $\{e_1,e_2,e_3\}$ possibly nonorthonormal there holds

$$\frac{\partial f}{\partial \tau} = \lambda_1\frac{\partial f}{\partial e_1} + \lambda_2\frac{\partial f}{\partial e_2} + \lambda_3\frac{\partial f}{\partial e_3} \tag{2.5}$$

on E, where the coefficients λ_1, λ_2, and λ_3 satisfy

$$\tau = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3. \tag{2.6}$$

Fix $\frac{\pi}{2} < \alpha < \pi$, $\nu \in S^2$ with $0 < \tau \cdot \nu < 1$. We will start by showing that if $\mathcal{H}^1(E \cap \Sigma^f) > 0$, then there exists a basis $\{e_1, e_2, e_3\}$ such that $e_3 \in S^2 \setminus \Pi_\nu$,

$$\{e_1, e_2\} \subset S^2 \cap \Pi_\nu, \qquad e_1 \cdot e_2 = \cos\alpha, \tag{2.7}$$

and identity (2.5) fails on a subset of $E \cap \Sigma^f$ of positive $\mathcal{H}^1$-measure.

Before giving the precise statement, it is convenient to introduce the following notation. Given ν and $\{e_1, e_2, e_3\}$ as before, we set

$$\sigma^{\nu,e_3} := \tau - \frac{\tau \cdot \nu}{e_3 \cdot \nu} e_3 = \pi_\nu(\tau) - \frac{\tau \cdot \nu}{e_3 \cdot \nu} \pi_\nu(e_3). \tag{2.8}$$

Note that the vector σ^{ν,e_3} belongs to $\Pi_\nu \cap \mathrm{span}\{\tau, e_3\}$.

Proposition 2.8. *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a Lipschitz continuous such that the partial derivative $\frac{\partial f}{\partial \sigma}$ exists $\mathcal{H}^1$-a.e. on E for every $\sigma \in S^2$. Assume also that $\mathcal{H}^1(\Sigma^f \cap E) > 0$, where E satisfies* (2.3)*. Then, given $\frac{\pi}{2} < \alpha < \pi$, $0 < \varepsilon < 1$, $0 < \delta < \cos^2 \frac{\alpha}{2}$, and $\nu \in S^2$ with $0 < \nu \cdot \tau < 1$, there exist $\{e_1, e_2\} \in S^2 \cap \Pi_\nu$ and $e_3 \in S^2 \setminus \Pi_\nu$ with $0 < |e_3 - \tau| \le \varepsilon$, such that* (2.7) *holds,*

$$(\sigma^{\nu,e_3} \cdot e_1)(\sigma^{\nu,e_3} \cdot e_2) \ge \delta |\sigma^{\nu,e_3}|^2, \tag{2.9}$$

and (2.5) *fails on some subset of E of positive $\mathcal{H}^1$-measure.*

Note that the quantity $\cos^2 \frac{\alpha}{2}$ (appearing in the statement of the proposition) represents the maximum possible value for δ in (2.9), which is achieved when σ^{ν,e_3} is parallel to $e_1 + e_2$.

Definition 2.9. *Let $0 < \varepsilon < 1$, $\frac{\pi}{2} < \alpha < \pi$, and $0 < \delta < \cos^2 \frac{\alpha}{2}$ be fixed. Given two bases $\{e_1, e_2\}$ and $\{\epsilon_1, \epsilon_2\}$ in $\Pi_\nu \cap S^2$, we say that they are* compatible *if $e_1 \cdot e_2 = \epsilon_1 \cdot \epsilon_2 = \cos\alpha$ and there exists $e_3 \in S^2$, with $0 < |e_3 - \tau| \le \varepsilon$, such that $\{e_1, e_2, e_3\}$ and $\{\epsilon_1, \epsilon_2, e_3\}$ are two bases of $\mathbb{R}^3$ satisfying* (2.9) *(the latter with e_i replaced by ϵ_i).*

Lemma 2.10. *Let ε, α, and δ as in Definition 2.9 and let $\mathcal{B}$ be a countable dense family of bases $\{\hat{e}_1, \hat{e}_2\}$ in $\Pi_\nu \cap S^2$ such that $\hat{e}_1 \cdot \hat{e}_2 = \cos\alpha$. Then, for any two given bases $\{e_1, e_2\}$ and $\{\epsilon_1, \epsilon_2\}$ in $\mathcal{B}$ there exists a finite chain of bases $\{\epsilon_1^j, \epsilon_2^j\}$ in $\mathcal{B}$, $j = 0, \dots, k$, such that $\{\epsilon_1^0, \epsilon_2^0\} = \{e_1, e_2\}$, $\{\epsilon_1^k, \epsilon_2^k\} = \{\epsilon_1, \epsilon_2\}$, and $\{\epsilon_1^{j-1}, \epsilon_2^{j-1}\}$ is compatible with $\{\epsilon_1^j, \epsilon_2^j\}$ according to Definition 2.9 for $j = 1, \dots, k$.*

Proof. First note that $\mathrm{span}\{\pm\sigma^{\nu,e_3} : e_3 \in S^2,\ 0 < |e_3 - \tau| < \varepsilon\} = \Pi_\nu$. Indeed, given $\eta \in \Pi_\nu$, it suffices to choose $e_3 \in S^2 \cap \mathrm{span}\{\eta, \tau\}$ with $0 < |e_3 - \tau| < \varepsilon$ to obtain that η and σ^{ν,e_3} are parallel.

Without loss of generality we will only consider oriented bases $\{e_1, e_2\}$ in $\Pi_\nu \cap S^2$ such that $\frac{e_1 \wedge e_2}{|e_1 \wedge e_2|} = \nu$. For any $e_3 \in S^2$ consider the set of all $e_1 \in \Pi_\nu \cap S^2$ such that there exists an oriented basis $\{e_1, e_2\}$ satisfying (2.7) and $(\sigma^{\nu,e_3} \cdot e_1)(\sigma^{\nu,e_3} \cdot e_2) > \delta |\sigma^{\nu,e_3}|^2$. Such a set consists of two open arcs on $\Pi_\nu \cap S^2$, symmetric with respect to the origin, whose length and distance from $\frac{\sigma^{\nu,e_3}}{|\sigma^{\nu,e_3}|}$ depend only on δ. By the remark made at the beginning of this proof, it follows that as e_3 varies in $S^2 \cap \{0 < |e_3 - \tau| < \varepsilon\}$, the corresponding intervals cover all of $\Pi_\nu \cap S^2$. By compactness we may extract a finite collection of open arcs $I_0, \dots, I_k$ on $\Pi_\nu \cap S^2$ such that $\cup_j I_j = \Pi_\nu \cap S^2$ and $I_j \cap I_{j+1} \ne \emptyset$. Moreover, by construction, if $\{\epsilon_1, \epsilon_2\}$ and $\{\hat{\epsilon}_1, \hat{\epsilon}_2\}$ satisfy (2.7) and both ϵ_1 and $\hat{\epsilon}_1$ belong to I_j for some $j \in \{0, \dots, k\}$, then the two bases are compatible. We are now in a position to conclude. Indeed, consider any pair $\{e_1, e_2\}$ and $\{\epsilon_1, \epsilon_2\}$ in $\mathcal{B}$ and assume that they are not compatible. Hence, there exist $0 \le i < j \le k$, with $j - i \ge 2$, such that $e_1 \in I_i$ and $\epsilon_1 \in I_j$. To conclude, it is enough to choose for $h = 0, \dots, j - i - 1$ a basis $\{\epsilon_1^h, \epsilon_2^h\} \in \mathcal{B}$ such that $\epsilon_1^h \in I_{i+h} \cap I_{i+h+1}$. This is possible due to the density of $\mathcal{B}$. □

Proof of Proposition 2.8. Assume by contradiction that (2.5) holds for all $\{e_1, e_2, e_3\}$ as in the statement and fix one such a triple. Note that $\nu = \frac{1}{e_3 \cdot \nu} e_3 - \frac{1}{e_3 \cdot \nu} \pi_\nu(e_3)$ and, in turn,

$$\tau = \frac{\tau \cdot \nu}{e_3 \cdot \nu} e_3 + \pi_\nu(\tau) - \frac{\tau \cdot \nu}{e_3 \cdot \nu} \pi_\nu(e_3). \tag{2.10}$$

Moreover, by (2.7) for any vector $v \in \Pi_\nu$ we have

$$v = \Big(v \cdot \Big(\frac{1}{\sin^2\alpha} e_1 - \frac{\cos\alpha}{\sin^2\alpha} e_2\Big)\Big) e_1 + \Big(v \cdot \Big(-\frac{\cos\alpha}{\sin^2\alpha} e_1 + \frac{1}{\sin^2\alpha} e_2\Big)\Big) e_2.$$

Hence, by (2.10), we may write

$$\tau = \lambda_1^{e_3} e_1 + \lambda_2^{e_3} e_2 + \frac{\tau \cdot \nu}{e_3 \cdot \nu} e_3, \tag{2.11}$$

where

$$\lambda_1^{e_3} := \sigma^{\nu,e_3} \cdot \Big(\frac{1}{\sin^2\alpha} e_1 - \frac{\cos\alpha}{\sin^2\alpha} e_2\Big) \quad \text{and} \quad \lambda_2^{e_3} := \sigma^{\nu,e_3} \cdot \Big(-\frac{\cos\alpha}{\sin^2\alpha} e_1 + \frac{1}{\sin^2\alpha} e_2\Big). \tag{2.12}$$

By our contradiction hypothesis, it follows that

$$\frac{\partial f}{\partial \tau}(w) = \lambda_1^{e_3} \frac{\partial f}{\partial e_1}(w) + \lambda_2^{e_3} \frac{\partial f}{\partial e_2}(w) + \frac{\tau \cdot \nu}{e_3 \cdot \nu} \frac{\partial f}{\partial e_3}(w) \tag{2.13}$$

for $\mathcal{H}^1$-a.e. w in E. Choose now different vectors ϵ_1 and ϵ_2 such that $\{\epsilon_1, \epsilon_2\} \subset \Pi_\nu \cap S^2$, $\epsilon_1 \cdot \epsilon_2 = \cos\alpha$, and

$$(\sigma^{\nu,e_3} \cdot \epsilon_1)(\sigma^{\nu,e_3} \cdot \epsilon_2) \geq \delta |\sigma^{\nu,e_3}|^2. \tag{2.14}$$

Then, by the same arguments, equality (2.13) holds with e_i and $\lambda_i^{e_3}$ replaced by ϵ_i and $\mu_i^{e_3}$, respectively, where

$$\mu_1^{e_3} := \sigma^{\nu,e_3} \cdot \Big(\frac{1}{\sin^2\alpha} \epsilon_1 - \frac{\cos\alpha}{\sin^2\alpha} \epsilon_2\Big) \quad \text{and} \quad \mu_2^{e_3} := \sigma^{\nu,e_3} \cdot \Big(-\frac{\cos\alpha}{\sin^2\alpha} \epsilon_1 + \frac{1}{\sin^2\alpha} \epsilon_2\Big). \tag{2.15}$$

By comparison we conclude that

$$\lambda_1^{e_3} \frac{\partial f}{\partial e_1}(w) + \lambda_2^{e_3} \frac{\partial f}{\partial e_2}(w) = \mu_1^{e_3} \frac{\partial f}{\partial \epsilon_1}(w) + \mu_2^{e_3} \frac{\partial f}{\partial \epsilon_2}(w) \tag{2.16}$$

for $\mathcal{H}^1$-a.e. $w \in E$.

Now let $\hat{e}_3 \in S^2$ be such that $0 < |\hat{e}_3 - \tau| \leq \varepsilon$, σ^{ν,e_3} and $\sigma^{\nu,\hat{e}_3} := \pi_\nu(\tau) - \frac{\tau \cdot \nu}{\hat{e}_3 \cdot \nu} \pi_\nu(\hat{e}_3)$ are linearly independent, and both (2.9) and (2.14) hold with σ^{ν,e_3} replaced by $\sigma^{\nu,\hat{e}_3}$. Then, arguing as for (2.16), we deduce

$$\lambda_1^{\hat{e}_3} \frac{\partial f}{\partial e_1}(w) + \lambda_2^{\hat{e}_3} \frac{\partial f}{\partial e_2}(w) = \mu_1^{\hat{e}_3} \frac{\partial f}{\partial \epsilon_1}(w) + \mu_2^{\hat{e}_3} \frac{\partial f}{\partial \epsilon_2}(w) \tag{2.17}$$

for $\mathcal{H}^1$-a.e. $w \in E$, where $\lambda_i^{\hat{e}_3}$ and $\mu_i^{\hat{e}_3}$ are defined as in (2.12) and (2.15), respectively, with σ^{ν,e_3} replaced by $\sigma^{\nu,\hat{e}_3}$. Due to the linear independence of the vectors σ^{ν,e_3} and $\sigma^{\nu,\hat{e}_3}$, it follows from (2.12), (2.15), (2.16), and (2.17), that

$$\begin{aligned} &\frac{\partial f}{\partial e_1}(w)\Big(\frac{1}{\sin^2\alpha} e_1 - \frac{\cos\alpha}{\sin^2\alpha} e_2\Big) + \frac{\partial f}{\partial e_2}(w)\Big(-\frac{\cos\alpha}{\sin^2\alpha} e_1 + \frac{1}{\sin^2\alpha} e_2\Big) \\ &= \frac{\partial f}{\partial \epsilon_1}(w)\Big(\frac{1}{\sin^2\alpha} \epsilon_1 - \frac{\cos\alpha}{\sin^2\alpha} \epsilon_2\Big) + \frac{\partial f}{\partial \epsilon_2}(w)\Big(-\frac{\cos\alpha}{\sin^2\alpha} \epsilon_1 + \frac{1}{\sin^2\alpha} \epsilon_2\Big) \end{aligned} \tag{2.18}$$

for $\mathcal{H}^1$-a.e. $w \in E$. Summarizing, we have proved that if two bases $\{e_1, e_2\}$ and $\{\epsilon_1, \epsilon_2\}$ of Π_ν are compatible according to Definition 2.9, then (2.18) holds.

Let now $\mathcal{B}$ be a countable dense family of bases $\{\hat{e}_1, \hat{e}_2\}$ in $\Pi_\nu \cap S^2$ such that $\hat{e}_1 \cdot \hat{e}_2 = \cos\alpha$. Then, by the above discussion and by Lemma 2.10, it follows that for $\mathcal{H}^1$-a.e. $w \in E$ the identity (2.18) holds for any pair of bases in $\mathcal{B}$. We deduce that there exists an invariant vector v such that for $\mathcal{H}^1$-a.e. $w \in E$

$$\frac{\partial f}{\partial \hat{e}_1}(w)\Big(\frac{1}{\sin^2\alpha} \hat{e}_1 - \frac{\cos\alpha}{\sin^2\alpha} \hat{e}_2\Big) + \frac{\partial f}{\partial \hat{e}_2}(w)\Big(-\frac{\cos\alpha}{\sin^2\alpha} \hat{e}_1 + \frac{1}{\sin^2\alpha} \hat{e}_2\Big) = v$$

and, in turn,

$$\frac{\partial f}{\partial \hat{e}_1}(w) = v \cdot \hat{e}_1 \tag{2.19}$$

for all $\{\hat e_1, \hat e_2\} \in \mathcal{B}$. Since the set of $\hat e_1$'s satisfying (2.19) is dense in $S^2 \cap \Pi_\nu$, it follows from Proposition 2.1 that f restricted to the hyperplane $w + \Pi_\nu$ is differentiable at w for $\mathcal{H}^1$-a.e. $w \in E$. By Proposition 2.6 this implies that f is differentiable at $\mathcal{H}^1$-a.e. $w \in E$ and contradicts the assumption $\mathcal{H}^1(E \cap \Sigma^f) > 0$. □

Proposition 2.11. *Let f and E be as in Proposition 2.8 and let $\{e_1, e_2, e_3\} \subset S^2$ be a basis such that $|\tau \cdot e_i| < 1$ for $i = 1, 2, 3$. Assume that for every $\sigma \in S^2$ the partial derivative $\frac{\partial f}{\partial \sigma}$ exists $\mathcal{H}^1$-a.e. on E. Assume also that (2.5) fails on some subset of E of positive $\mathcal{H}^1$-measure. Then the same happens also with respect to any orthonormal coordinate system $\{\hat e_1, \hat e_2, \hat e_3\}$ such that $\hat e_i \in \mathrm{span}\,\{\tau, e_i\}$ and $\hat e_i \cdot \tau \neq 0$, $i = 1, 2, 3$.*

Proof. From the assumptions we may find a subset $F \subset E$ of positive $\mathcal{H}^1$-measure, a function η, and a constant $\eta_0 > 0$ such that

$$\frac{\partial f}{\partial \tau}(w) = \lambda_1 \frac{\partial f}{\partial e_1}(w) + \lambda_2 \frac{\partial f}{\partial e_2}(w) + \lambda_3 \frac{\partial f}{\partial e_3}(w) + \eta(w) \quad \text{for all } w \in F, \tag{2.20}$$

where the coefficients λ_i satisfy (2.6) and $|\eta(w)| \geq \eta_0 > 0$ for all $w \in F$.

Let $\{\hat e_1, \hat e_2, \hat e_3\}$ be an orthonormal basis as in the statement. Clearly we can write $\hat e_i = c_1^{(i)} e_i + c_2^{(i)} \tau$ for some $c_1^{(i)}$, $c_2^{(i)} \in \mathbb{R}$, with $c_1^{(i)} \neq 0$. Using the fact that by assumption $\frac{\partial f}{\partial e_i}$ exists $\mathcal{H}^1$-a.e. on F, it follows from Proposition 2.6 (applied to f restricted to $f(0) + \mathrm{span}\{e_i, \tau\}$) that the restriction of f to the affine plane $f(0) + \mathrm{span}\{e_i, \tau\}$, $i = 1, 2, 3$, is differentiable at $\mathcal{H}^1$-a.e. $w \in F$. For all such w's we have

$$\frac{\partial f}{\partial \hat e_i}(w) = c_1^{(i)} \frac{\partial f}{\partial e_i}(w) + c_2^{(i)} \frac{\partial f}{\partial \tau}(w). \tag{2.21}$$

Combining (2.20) and (2.21), after some elementary algebraic manipulations we arrive at

$$\frac{\partial f}{\partial \tau}(w) = \sum_{i=1}^{3} \eta_1 \frac{\lambda_i}{c_1^{(i)}} \frac{\partial f}{\partial \hat e_i}(w) + \eta_1 \eta(w) \tag{2.22}$$

for $\mathcal{H}^1$-a.e. $w \in F$, where[2]

$$\eta_1 := \prod_{i=1}^{3} \Big(1 + \frac{c_2^{(i)}}{c_1^{(i)}} \lambda_i\Big)^{-1} \neq 0.$$

The conclusion then follows from (2.22) once we note that $\tau = \sum_{i=1}^3 \eta_1 \frac{\lambda_i}{c_1^{(i)}} \hat e_i$. □

2.4. **Proof of the main theorems.** We are now in a position to prove Theorem 1.1.

Proof of Theorem 1.1. We split the proof into several steps.

Step 1. The case $k = 1$, $d = 3$, and E contained in a straight line. Let E satisfy (2.3) and assume by contradiction that $\mathcal{H}^1(E \cap \Sigma^f) > 0$. By Corollary 2.4 it suffices to show that under these circumstances there exists an orthonormal basis $\{\hat e_1, \hat e_2, \hat e_3\}$ for which (2.4) fails on some subset of E of positive $\mathcal{H}^1$-measure..

Recall that by Remark 2.5 for every $\sigma \in S^2$ the partial derivative $\frac{\partial f}{\partial \sigma}$ exists $\mathcal{H}^1$-a.e. on E. For every $\varepsilon > 0$ and $\nu^\varepsilon \in S^2$, with $0 < |\nu^\varepsilon - \tau| \leq \varepsilon$, by Proposition 2.8 there exists a basis $\{e_1^\varepsilon, e_2^\varepsilon, e_3^\varepsilon\} \subset S^2$ such that such that e_1^ε, $e_2^\varepsilon \in \Pi_{\nu^\varepsilon}$, $e_1^\varepsilon \cdot e_2^\varepsilon = \cos\alpha$, $0 < |e_3^\varepsilon - \tau| \leq \varepsilon$,

$$(\sigma^{\nu^\varepsilon, e_3^\varepsilon} \cdot e_1^\varepsilon)(\sigma^{\nu^\varepsilon, e_3^\varepsilon} \cdot e_2^\varepsilon) \geq \delta |\sigma^{\nu^\varepsilon, e_3^\varepsilon}|^2, \tag{2.23}$$

with $\sigma^{\nu^\varepsilon, e_3^\varepsilon}$ defined as in (2.8), and (2.5) fails with respect to $\{e_1^\varepsilon, e_2^\varepsilon, e_3^\varepsilon\}$ on a subset of E of positive $\mathcal{H}^1$-measure. Let $\varepsilon_n \searrow 0$ and for $i = 1, 2$ let $\hat e_i^{\varepsilon_n}$ be uniquely determined by

$$\hat e_i^{\varepsilon_n} \in S^2 \cap \Pi_{e_3^{\varepsilon_n}} \cap \mathrm{span}\{e_i^{\varepsilon_n}, \tau\} \quad \text{and} \quad \hat e_i^{\varepsilon_n} \cdot e_i^{\varepsilon_n} > 0. \tag{2.24}$$

[2]It is easy to check that the definition of η_1 is well-posed thanks to the linear independence of e_1, e_2, and e_3.

Since $|\nu^{\varepsilon_n} - e_3^{\varepsilon_n}| \to 0$, it follows from (2.24) that $|\hat{e}_i^{\varepsilon_n} - e_i^{\varepsilon_n}| \to 0$ and, in turn,

$$\hat{e}_1^{\varepsilon_n} \cdot \hat{e}_2^{\varepsilon_n} \to \cos\alpha. \tag{2.25}$$

Without loss of generality we may also assume that there exist a plane P and unit vectors σ and $\bar{\sigma}$ such that $\mathrm{span}\{\tau, e_3^{\varepsilon_n}\} \to P$, $\frac{\pi_{e_3^{\varepsilon_n}}(\tau)}{|\pi_{e_3^{\varepsilon_n}}(\tau)|} \to \sigma$, and $\frac{\sigma^{\nu^{\varepsilon_n}, e_3^{\varepsilon_n}}}{|\sigma^{\nu^{\varepsilon_n}, e_3^{\varepsilon_n}}|} \to \bar{\sigma}$. As $\frac{\pi_{e_3^{\varepsilon_n}}(\tau)}{|\pi_{e_3^{\varepsilon_n}}(\tau)|} \in \Pi_{e_3^{\varepsilon_n}} \cap \mathrm{span}\{\tau, e_3^{\varepsilon_n}\}$, while $\frac{\sigma^{\nu^{\varepsilon_n}, e_3^{\varepsilon_n}}}{|\sigma^{\nu^{\varepsilon_n}, e_3^{\varepsilon_n}}|} \in \Pi_{\nu^{\varepsilon_n}} \cap \mathrm{span}\{\tau, e_3^{\varepsilon_n}\}$ by (2.8), since $e_3^{\varepsilon_n}$ and ν^{ε_n} converge to τ, we deduce that both σ and $\bar{\sigma}$ belong to $P \cap \Pi_\tau$; that is, either $\sigma = \bar{\sigma}$ or $\sigma = -\bar{\sigma}$. In either case, by (2.23) we have

$$\lim_{n\to\infty} \Big(\frac{\pi_{e_3^{\varepsilon_n}}(\tau)}{|\pi_{e_3^{\varepsilon_n}}(\tau)|} \cdot \hat{e}_1^{\varepsilon_n} \Big) \Big(\frac{\pi_{e_3^{\varepsilon_n}}(\tau)}{|\pi_{e_3^{\varepsilon_n}}(\tau)|} \cdot \hat{e}_2^{\varepsilon_n} \Big) \geq \delta. \tag{2.26}$$

Hence, by (2.24), (2.25), and (2.26) we can fix a basis, denoted for simplicity $\{e_1, e_2, e_2\}$, such that (2.5) fails with respect to $\{e_1, e_2, e_2\}$ on a subset of E of positive $\mathcal{H}^1$-measure and

$$\hat{e}_1 \cdot \hat{e}_2 < 0\,, \qquad (\pi_{e_3}(\tau) \cdot \hat{e}_1)(\pi_{e_3}(\tau) \cdot \hat{e}_2) > 0\,, \tag{2.27}$$

where $\hat{e}_i$ is uniquely determined by

$$\hat{e}_i \in S^2 \cap \Pi_{e_3} \cap \mathrm{span}\{e_i, \tau\} \quad \text{and} \quad \hat{e}_i \cdot e_i > 0.$$

By Proposition 2.11 the same happens with respect to any orthonormal basis $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ such that $\epsilon_i \in \mathrm{span}\{e_i, \tau\}$ and $\epsilon_i \cdot \tau \neq 0$, $i = 1, 2, 3$. Hence, it remains to show that such a basis exists.

Without loss of generality we may assume that $\tau \cdot e_3 > 0$ and, by (2.27), that

$$\pi_{e_3}(\tau) \cdot \hat{e}_i > 0 \quad i = 1, 2. \tag{2.28}$$

Let $\mu \in S^2 \cap \Pi_{e_3}$ be orthogonal to $\mathrm{span}\{e_3, \tau\}$ and for $s \in [0,1]$ set $v(s) := \pi_{e_3}(\tau) + s(\tau \cdot e_3)e_3$. For $i = 1, 2$ let $e_i(s)$ be uniquely determined by

$$e_i(s) \in \mathrm{span}\{\mu, v(s)\} \cap \mathrm{span}\{e_i, \tau\}, \quad \pi_{e_3}(\tau) \cdot e_i(s) > 0\,. \tag{2.29}$$

The map $s \mapsto e_i(s)$ turns out to be continuous. Moreover, using (2.28) and (2.29) we deduce $e_i(0) = \hat{e}_i$ and $e_i(1) = \tau$ for $i = 1, 2$. It follows, in particular, that $e_1(0) \cdot e_2(0) = \hat{e}_1 \cdot \hat{e}_2 < 0$ thanks to (2.27), while $e_1(1) \cdot e_2(1) = 1$. Hence, by continuity $e_1(\bar{s}) \cdot e_2(\bar{s}) = 0$ for some $\bar{s} \in (0,1)$. Set $\epsilon_i = e_i(\bar{s})$ for $i = 1, 2$ and let $\epsilon_3 \in S^2$ be orthogonal to $\mathrm{span}\{\epsilon_1, \epsilon_2\}$. By constrcution ϵ_1 and ϵ_2 are orthogonal and belong to $\mathrm{span}\{e_1, \tau\}$ and $\mathrm{span}\{e_2, \tau\}$, respectively. Recalling that μ is orthogonal to $\mathrm{span}\{e_3, \tau\}$ we have

$$\epsilon_3 \in (\mathrm{span}\{\epsilon_1, \epsilon_2\})^\perp = (\mathrm{span}\{\mu, v(\bar{s})\})^\perp \subset \mathrm{span}\{e_3, \tau\}$$

and we conclude that $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ is an orthonormal basis with all the required properties.

Step 2. The case $k = 1$, $d \geq 3$, and E contained in a straight line. We proceed by induction on the dimension d. Assume that the theorem is true for some $d = j \geq 3$, under the additional assumption $E \subset \mathrm{span}\{\tau\}$. We will show that it holds also for $d = j + 1$, under the same additional hypothesis. To this aim, choose a dense sequence $\{P_n\}$ of $(d-2)$-dimensional subspaces of the hyperplane Π_τ. By Corollary 2.4, the restriction of f to each hyperplane $\mathrm{span}\{\tau\} + P_n$ satisfies the assumptions of the theorem and hence is differentiable at $\mathcal{H}^1$-a.e. $w \in E$ by the inductive hypothesis. In particular, for $\mathcal{H}^1$-a.e. $w \in E$ the restriction of f to $w + P_n$ is differentiable at w for every n. In turn, by Proposition 2.3, for $\mathcal{H}^1$-a.e. $w \in E$ the restriction of f to any $(d-2)$-dimensional affine subspace of $w + \Pi_\tau$ through w is differentiable at w. The conclusion then follows from Proposition 2.1 (applied to f restricted to $w + \Pi_\tau$) and Proposition 2.6.

Step 3. The case $k = 1$, $d \geq 3$, and E general. We argue by contradiction by assuming that there exists a regular curve $\gamma : [0,1] \to \mathbb{R}^d$ of class C^1 such that $\mathcal{H}^1(\gamma([0,1]) \cap E \cap \Sigma^f) > 0$, where Σ^f denotes, as usual, the singular set (1.2). Let w_0 be a point of positive $\mathcal{H}^1$-density in $\gamma([0,1]) \cap E \cap \Sigma^f$. We claim that for ε small enough there exists a local diffeomorphism $\Phi : B(w_0;\varepsilon) \to \mathbb{R}^d$ such that

$$\Phi(\gamma([0,1]) \cap B(w_0;\varepsilon)) \subset \operatorname{span}\{\tau\} \quad \text{and} \quad D\Phi(w) \in SO^+(d) \text{ for } w \in \gamma([0,1]) \cap B(w_0;\varepsilon), \tag{2.30}$$

where $SO^+(d)$ denotes the set of orthogonal $(d \times d)$-matrices with determinant equal to 1. To this aim fix $\tau \in S^{d-1}$. Notice that if ε is sufficiently small then we may find two continuous maps $t : \gamma[0,1] \cap \overline{B(w_0;\varepsilon)} \to S^{d-1}$ and $Q : \gamma[0,1] \cap \overline{B(w_0;\varepsilon)} \to SO(d)$ such that $t(w)$ is tangent to γ at w and $Q(w)[t(w)] = \tau$ for all $w \in \gamma[0,1] \cap \overline{B(w_0;\varepsilon)}$. By integrating Q along γ we define a C^1-map $\Phi : \gamma[0,1] \cap \overline{B(w_0;\varepsilon)} \to \mathbb{R}^d$ whose tangential gradient coincides with τ. We may now apply the Whitney Extension Theorem (see for instance [5, Section 6.5]) to extend Φ to a C^1-map $\Phi : B(w_0;\varepsilon) \to \mathbb{R}^d$ whose gradient coincides with Q along γ. The claim is proved.

In turn, there exists a Lipschitz function $g : \mathbb{R}^d \to \mathbb{R}$ that coincides with $f \circ \Phi^{-1}$ in $\Phi(B(w_0;\varepsilon))$. Clearly, $\mathcal{H}^1(\Sigma^g \cap \Phi(\gamma([0,1]) \cap B(w_0;\varepsilon))) > 0$. Recalling (2.30), it follows from Step 2 that there exists an orthonormal basis $\{e_1, \dots e_d\}$ such that (1.1) (with f replaced by g) fails with respect to $\{e_1, \dots e_d\}$ on a subset of $\Phi(\gamma([0,1]) \cap B(w_0;\varepsilon))$ of positive $\mathcal{H}^1$-measure. In terms of the original function f, this is equivalent to saying that the set

$$\mathcal{H}^1\Big(\Big\{w \in \gamma([0,1]) \cap B(w_0;\varepsilon) : \frac{\partial f}{\partial t(w)}(w) \neq \sum_{i=1}^d (t(w)\cdot e_i(w)) \frac{\partial f}{\partial e_i(w)}(w)\Big\}\Big) > 0\,,$$

where $e_i(w) := D\Phi^{-1}(\Phi(w))[e_i]$. Note that by (2.30) $\{e_1(w), \dots e_d(w)\}$ is an orthonormal basis for all $w \in \gamma([0,1]) \cap B(w_0;\varepsilon)$. We may now find $\delta > 0$ and $w_1 \in \gamma([0,1]) \cap B(w_0;\varepsilon)$ such that

$$E^\delta := \Big\{w \in \gamma([0,1]) \cap B(w_0;\varepsilon) : \Big|\frac{\partial f}{\partial t(w)}(w) - \sum_{i=1}^d (t(w)\cdot e_i(w)) \frac{\partial f}{\partial e_i(w)}(w)\Big| > \delta\Big\} \tag{2.31}$$

has positive $\mathcal{H}^1$-measure and w_1 has positive density in E^δ. To conclude the proof it suffices to show that

$$\frac{\partial f}{\partial t(w)}(w) \neq \sum_{i=1}^d (t(w)\cdot e_i(w_1)) \frac{\partial f}{\partial e_i(w_1)}(w)$$

in $E^\delta \cap B(w_1;\eta)$ for η small enough. But this follows from (2.31) and the fact that

$$\sum_{i=1}^d (t(w)\cdot e_i(w)) \frac{\partial f}{\partial e_i(w)}(w) - \sum_{i=1}^d (t(w)\cdot e_i(w_1)) \frac{\partial f}{\partial e_i(w_1)}(w) \to 0 \qquad \text{as } w \to w_1,$$

which, in turn, is a consequence of the continuity of $w \mapsto e_i(w)$ and of (2.2).

Step 4. Conclusion. We consider here the general case $k \geq 1$. Let $M \subset \mathbb{R}^d$ be a k-dimensional manifold such that $\mathcal{H}^k(M \cap E) > 0$. It suffices to show that for every $w_0 \in M \cap E$ the function f is differentiable $\mathcal{H}^k$-a.e. in $M \cap B(w_0;\varepsilon)$ for ε small enough. Fix any $w_0 \in M \cap E$ and consider a local regular parametrization $\psi : D \subset \mathbb{R}^k \to M$ of class C^1 such that $M \cap B(w_0;\varepsilon) \subset \psi(D)$, where D is an open k-dimensional set and $\psi(0) = w_0$. Such a local parametrization exists if ε is small enough. Without loss of generality assume that

$$D = B_1(0;r) \times B_{k-1}(0;r)$$

and write $x = (x_1, x') \in \mathbb{R} \times \mathbb{R}^{k-1}$. For every $x' \in B_{k-1}(0;r)$ let $E_{x'}$ denote the $\mathcal{H}^1$-rectifiable set obtained by intersecting E with the support of the curve $\psi(\cdot, x')$. Recalling (1.1) and the fact that $\mathcal{S}$ is countable, by Fubini's Theorem there exists a set $\mathcal{M} \subset B_{k-1}(0;r)$ with $\mathcal{L}^{k-1}(\mathcal{M}) = 0$ such that for all $x' \in B_{k-1}(0;r) \setminus \mathcal{M}$ and for every coordinate system $\{e_1, \dots, e_d\} \in \mathcal{S}$ we have

$$\pi_w^E\big(\nabla^{(e_1 \dots e_d)} f(w)\big) = g(w) \quad \text{for } \mathcal{H}^1\text{-a.e. } w \in E_{x'}.$$

This, in turn, implies that for all $x' \in B_{k-1}(0;r) \setminus \mathcal{M}$ and for all $\{e_1, \dots, e_d\} \in \mathcal{S}$

$$\pi_w^{E_{x'}}\big(\nabla^{(e_1 \dots e_d)} f(w)\big) = \pi_w^{E_{x'}}(g(w)) \quad \text{for } \mathcal{H}^1\text{-a.e. } w \in E_{x'}.$$

By Step 3 it follows that for all $x' \in B_{k-1}(0;r) \setminus \mathcal{M}$ the function f is differentiable $\mathcal{H}^1$-a.e. on $E_{x'}$. By Fubini's Theorem again and the regularity of ψ, the last statement is equivalent to saying that f is differentiable $\mathcal{H}^k$-a.e. in $\psi(D) \cap E$. □

We finally prove Theorems 1.2 and 1.3 that are now an easy corollary of Theorem 1.1.

Proof of Theorems 1.2 and 1.3. Note that Theorem 1.2 is a particular case of Theorem 1.3, since $A_1(\mathbb{R}^N;\mathbb{R}^d) = W^{1,1}_{\rm loc}(\mathbb{R}^N;\mathbb{R}^d)$. As remarked in the introduction, we only have to prove the necessity part. Assume by contradiction that Σ^f is not purely $\mathcal{H}^k$-unrectifiable, $1 \leq k \leq \min\{\mathbb{N}, d\}$. Then we may find a regular parametrization $\psi : [0,1]^k \to \mathbb{R}^d$ of class C^1 such that $\mathcal{H}^k(\psi([0,1]^k) \cap \Sigma^f) > 0$. By Theorem 1.1 there exists an orthonormal basis $\{e_1, \dots, e_d\} \in \mathcal{S}$ for which (1.1) fails on a subset $F \subset \psi([0,1]^k) \cap \Sigma^f$ of positive $\mathcal{H}^k$-measure. Write every $x \in \mathbb{R}^N$ as $x = (x_1, x')$, where $x_1 \in \mathbb{R}^k$ and $x' \in \mathbb{R}^{N-k}$. It is now easy to check that if $u \in A_k(\mathbb{R}^N;\mathbb{R}^d)$ satisfies $u(x_1, x') = \psi(x_1)$ for $(x_1, x') \in (0,1)^N$, then the chain rule with respect to the basis $\{e_1, \dots, e_d\}$ fails for $f \circ u$ in $\psi^{-1}(F) \times (0,1)^{N-k}$, which has positive $\mathcal{L}^N$-measure, thanks to the regularity of ψ. The openness of the set of othonormal frames for which (1.3) fails follows immediately from (2.2). □

Acknowledgments

The author wishes to thank G. Leoni for interesting discussions on the subject of this paper.

References

[1] Ambrosio L. and G. Dal Maso, *A general chain rule for distributional derivatives*, Proc. Amer. Math. Soc. 108 (1990), pp. 691–702.

[2] Ambrosio L., Fusco N. and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, Inc. New York, 2000.

[3] Bessis D. N. and Clarke F. H., *Partial subdifferentials, derivates and Rademacher's theorem.* Trans. Amer. Math. Soc. 351 (1999), no. 7, 2899–2926.

[4] De Cicco V., Leoni G, *A chain rule in $L^1(\mathrm{div};\Omega)$ and its applications to lower semicontinuity*, Calc. Var. Partial Differential Equations 19 (2004), pp. 23–51.

[5] Evans L. C., Gariepy R. F., *Measure theory and fine properties of functions*, Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[6] Federer H., *Geometric measure theory*, Springer-Verlag, 1969.

[7] G. Leoni and M. Morini, *Necessary and sufficient conditions for the chain rule in $W^{1,1}_{loc}(\mathbb{R}^N,\mathbb{R}^d)$ and $BV_{loc}(\mathbb{R}^N,\mathbb{R}^d)$*, J. Eur. Math. Soc. (JEMS) 9 (2007), pp. 219–252.

[8] Marcus M. and V. J. Mizel, *Absolute continuity on tracks and mappings of Sobolev spaces*, Arch. rat. Mech. Anal. 45 (1972), pp. 294–320.

[9] Marcus M. and V. J. Mizel, *Continuity of certain Nemitsky operators on Sobolev spaces and the chain rule*, J. Analyse Math. 28 (1975), pp. 303–334.

[10] Marcus M. and V. J. Mizel, *Complete characterization of functions which act, via superposition, on Sobolev spaces*, Trans. Amer. Math. Soc. 251 (1979), pp. 187–218.

[11] Serrin J. and D. Varberg, *A general chain rule for derivatives and the change of variables formula for the Lebesgue integral*, Am. Math. Mon. 76 (1969), pp. 514–520.

(Massimiliano Morini) SISSA, Via Beirut 2, 34014 Trieste, Italy
E-mail address, Massimiliano Morini: `morini@sissa.it`