RESEARCH

Real Analysis Exchange
Vol. 40(1), 2014/2015, pp. 81–98

Ettore Minguzzi,* Dipartimento di Matematica e Informatica "U. Dini", Università degli Studi di Firenze, Via S. Marta 3, I-50139 Firenze, Italy.
email: ettore.minguzzi@unifi.it

THE EQUALITY OF MIXED PARTIAL DERIVATIVES UNDER WEAK DIFFERENTIABILITY CONDITIONS

Abstract

We review and develop two little known results on the equality of mixed partial derivatives, which can be considered the best results so far available in their respective domains. The former, due to Mikusiński and his school, deals with equality at a given point, while the latter, due to Tolstov, concerns equality almost everywhere. Applications to differential geometry and General Relativity are commented.

1 Introduction

As it is well known, under reasonable conditions the mixed partial derivatives of a real function coincide. This result has a long history, and several distinguished scholars provided proofs, including Euler and Clairaut. However, according to Lindelöf, none of those proofs was free of errors or tacit assumptions, so that historians give credit for the first correct proof to H. A. Schwarz; see [1] for a nice historical account.

Actually, the first correct proof of the equality of mixed partial derivatives was obtained by Cauchy, who improved and amended a previous proof by Lagrange. However, they assumed the existence and continuity of the derivatives $\partial_1^2 f$, $\partial_2^2 f$. Schwarz removed this assumption and showed also that the continuity of $\partial_1 \partial_2 f$ could be obtained from the other hypothesis. Let $O = (a, b) \times (c, d) \subset \mathbb{R}^2$. He proved [2]:

Mathematical Reviews subject classification: Primary: 26B05, 26B30; Secondary: 26A16, 97I40, 83C99
Key words: Schwarz's theorem, Fubini's theorem, Lipschitz connections
Received by the editors January 4, 2014
Communicated by: Miroslav Zeleny
*Research partially supported by GNFM of INDAM

1. Let $f \in C^1(O, \mathbb{R})$ and suppose that $\partial_2\partial_1 f$ exists and belongs to $C(O, \mathbb{R})$. Then $\partial_1\partial_2 f$ exists and $\partial_1\partial_2 f = \partial_2\partial_1 f$.

It is natural to ask whether the assumptions can be weakened. Stronger versions can be found in the first published studies of this problem. For instance, Dini in his "Lezioni di Analisi Infinitesimale" [3, p. 164] does not assume $f \in C^1(O, \mathbb{R})$, but demands just the existence of the partial derivatives and the continuity of $\partial_2 f$ in y.[1]

The strongest result in this direction seems to have been obtained by Peano who removed the assumption on the continuity of $\partial_2 f$ from Dini's version [4]. Peano's version can be found in Rudin [5].

2. Let $f : O \to \mathbb{R}$. Suppose that $\partial_1 f$, $\partial_2 f$ and $\partial_2\partial_1 f$ exist on O and that the latter is continuous at (x_0, y_0). Then $\partial_1\partial_2 f(x_0, y_0) = \partial_2\partial_1 f(x_0, y_0)$.

The continuity of f is not assumed, and there are indeed discontinuous functions which admit everywhere partial derivatives at any order [6].

Many authors tried to weaken the conditions of Schwarz's theorem in other directions. Young proved the following result (see Apostol [7, Theor. 12.12]):

3. Let $f : O \to \mathbb{R}$. If both partial derivatives $\partial_1 f$ and $\partial_2 f$ exist in neighborhood of $(x_0, y_0) \in O$ and if both are differentiable at (x_0, y_0), then $\partial_2\partial_1 f(x_0, y_0) = \partial_1\partial_2 f(x_0, y_0)$.

We observe that these assumptions imply that $\partial_1 f$ and $\partial_2 f$, being continuous at (x_0, y_0), are bounded in a neighborhood of this point. Thus f is Lipschitz and hence continuous in such neighborhood. As with the Lagrange-Cauchy version, Young's result assumes the existence of both $\partial_1^2 f(x_0, y_0)$ and $\partial_2^2 f(x_0, y_0)$.

We are now going to prove a result which improves Peano's. It is based on the concept of strong differentiation, also introduced by himself in [8, 9]. It seems that he did not realize the usefulness of strong differentiation for the problem of the equality of mixed derivatives, possibly because he investigated the latter problem before the introduction of this derivative.

Remark 1. Recently the notion of strong differentiation has received renewed attention since it has been proved that the exponential map of Lipschitz connections or sprays over $C^{2,1}$ manifolds is strongly differentiable at the origin [10]. This fact implies that the exponential map is a Lipeomorphism near the

[1] In Dini's book $\partial^2 f/\partial x\partial y$ means $\partial_2\partial_1 f$. We stress that in the terminology of this article a necessary condition for a limit, such as a partial derivative, to exist will be its finiteness. That is, we do not tacitly use the extended real line, as some other authors do. This convention allows us to write just "exists" in place of "exists and is finite".

origin. Thus this notion proves important to do differential geometry under weak differentiability conditions.

Definition 2. A function $f\colon B \to \mathbb{R}^c$, $B \subset \mathbb{R}^a \times \mathbb{R}^b$, $(x,z) \mapsto f(x,z)$, is said to be *partially strongly differentiable* with respect to x at $(x_0, z_0) \in \bar{B}$, with differential $\partial_1 f(x_0, z_0)$, if for every $\epsilon > 0$, there is a $\delta > 0$ such that for every x_1, x_2, z such that $\|x_1 - x_0\| < \delta$, $\|x_2 - x_0\| < \delta$, $\|z - z_0\| < \delta$, $(x_1, z) \in B$, $(x_2, z) \in B$,

$$\|f(x_2, z) - f(x_1, z) - \partial_1 f(x_0, y_0)(x_2 - x_1)\| \le \epsilon \|x_2 - x_1\|.$$

If f does not depend on z then $\partial_1 f$ is called *differential* and is denoted df.

Clearly, if a function is strongly differentiable then it is differentiable, and the strong differential coincides with the differential. Some interesting properties are [8, 11, 12]:

(i) If f is strongly differentiable at p, then its satisfies a Lipschitz condition in a neighborhood of p.

(ii) If f is differentiable in a neighborhood of p and the differential is continuous at p, then it is strongly differentiable at p. Conversely, if f is strongly differentiable at p and the differential exists in a neighborhood of p, then the differential is continuous at p. A similar version for partial strong differentiation holds (this point is an easy consequence of the mean value theorem).

(iii) If f is strongly differentiable over a subset $A \subset E$, then the strong differential is continuous over A with respect to the induced topology.

In particular, a function is strongly differentiable in an open set O if and only if it is continuously differentiable on O.

The concept of strong differentiation has some advantages over that of ordinary (Frechet) differentiation. In particular, it allows us to obtain simpler and stronger results through shorter proofs. It serves better the intuition and at the elementary level could possibly replace the usual differentiation in elementary textbooks on analysis. Indeed, it extends the the range of applicability of some key results in analysis by removing some continuity assumptions on derivatives. For instance:

(iv) A function which is partially strongly differentiable with respect to all its variables at a point p is also totally strongly differentiable at that point p [12].

(v) If a function $f: \mathbb{R} \to \mathbb{R}$ has positive strong derivative at a point, then it is increasing in a neighborhood of that point. More generally, if a function $f : \mathbb{R}^n \to \mathbb{R}^n$ has invertible strong differential at a point p, then it is injective in a neighborhood of p and the inverse is strongly differentiable at $f(p)$ with $df^{-1}(f(p)) = (df(p))^{-1}$ (Leach's inverse function theorem [13]).

Let us observe that the usual assumptions that make the corresponding results hold for ordinary differentiation imply that the differential is continuous in a neighborhood of p. As observed above, these assumptions serve essentially to assume strong differentiability in a neighborhood without naming it. The point of using strong differentiability is that strong differentiability at a point suffices.

Peano's theorem on the equality of mixed partial derivatives at (x_0, y_0) demands the existence of $\partial_2\partial_1 f$ in a neighborhood of (x_0, y_0) and its continuity there. By (ii) above, $\partial_1 f$ is partially strongly differentiable with respect to y at (x_0, y_0). Thus one can ask whether the previous conditions can be replaced by partial strong differentiability. The answer is affirmative. The author reobtained the next theorem unaware of a previous result by Mikusiński [14, 15, 16], subsequently generalized to Banach spaces by Skórnik [17]. Mikusiński calls "full derivative" the Peano's strong derivative and does not give references so that it was quite hard to spot his important work [16].

He reobtains first results due to Peano and more advanced results such as Leach's inverse function theorem, but he also obtains new results on the role of strong differentiation in integration theory. In [16, Ch. 12] he provides the best and most complete introduction to strong derivatives to date. The fact that he published those results in Polish [15], and in some sections in a book devoted to quite different problems, did not help to spread knowledge of his important contributions. For completeness, we include the next proof, as it is different from Mikusiński's and has weaker assumptions.

Theorem 3. *Let $f: O \to \mathbb{R}$. Suppose that the partial derivative $\partial_1 f$ exists on O and that it is partially strongly differentiable with respect to y at (x_0, y_0). Then, denoting with $A \subset O$ the subset where $\partial_2 f$ exists, provided $(x_0, y_0) \in \bar{A}$, $\partial_2 f (:= \partial_2 f|_A)$ is partially strongly differentiable with respect to x at (x_0, y_0) and $\partial_1\partial_2 f(x_0, y_0) = \partial_2\partial_1 f(x_0, y_0)$.*

We stress that while the assumptions are weaker than Peano's, the conclusion is stronger. For instance, the previous theorem implies that if $\partial_1\partial_2 f(\cdot, y_0)$ exists in a neighborhood of x_0, then it is continuous at x_0.

Esser and Shisha [11, 12] construct a simple function $h(y)$ defined on an open set of $y = 0$ which is not everywhere differentiable on any neighborhood of 0, but which is strongly differentiable at 0. Then $f(x, y) = xh(y)$ satisfies the assumptions of our theorem, but not those of Peano's.

PROOF. Let $\epsilon > 0$. Since $\partial_1 f$ is partially strongly differentiable with respect to y at (x_0, y_0), there is $\delta(\epsilon) > 0$ such that for every $\tilde{x} \in (a, b)$, $\tilde{y}_1, \tilde{y}_2 \in (c, d)$, $|\tilde{x} - x_0| < \delta$, $|\tilde{y}_1 - y_0| < \delta$, $|\tilde{y}_2 - y_0| < \delta$, we have

$$|\partial_1 f(\tilde{x}, \tilde{y}_2) - \partial_1 f(\tilde{x}, \tilde{y}_1) - \partial_2\partial_1 f(x_0, y_0)(\tilde{y}_2 - \tilde{y}_1)| \le \epsilon|\tilde{y}_2 - \tilde{y}_1|. \tag{1}$$

Given $\epsilon > 0$, let $\delta(\epsilon) > 0$ be as above. Let $x_1, x_2 \in (a, b)$ be such that $|x_1 - x_0| < \delta$, $|x_2 - x_0| < \delta$, and let $y \in (c, d)$ be such that $|y - y_0| < \delta$. Furthermore, let them be such that $(x_1, y) \in A$, $(x_2, y) \in A$.

Let $y_1, y_2 \in (c, d)$, $y_1 \neq y_2$, be arbitrary and such that $|y_1 - y_0| < \delta$, $|y_2 - y_0| < \delta$. Let $u(t) := f(t, y_2) - f(t, y_1)$, then by the existence of $\partial_1 f$ and by the mean value theorem there is $x \in (a, b)$, $|x - x_0| < \delta$, such that

$$\partial_1 u(x)(x_2 - x_1) = u(x_2) - u(x_1).$$

Equation (1) holds for these values for x, y_1, y_2. Thus

$$\begin{aligned}|f(x_2, y_2) - f(x_2, y_1) - f(x_1, y_2) + f(x_1, y_1) - \partial_2\partial_1 f(x_0, y_0)(y_2 - y_1)(x_2 - x_1)| \\ \le \epsilon|x_2 - x_1|\,|y_2 - y_1|.\end{aligned}$$

Dividing by $|y_2 - y_1|$, setting $y_1 = y$ and taking the limit $y_2 \to y$, we obtain

$$|\partial_2 f(x_2, y) - \partial_2 f(x_1, y) - \partial_2\partial_1 f(x_0, y_0)(x_2 - x_1)| \le \epsilon|x_2 - x_1|,$$

which means that $\partial_2 f$ has partial strong differential with respect to x at (x_0, y_0) given by $\partial_2\partial_1 f(x_0, y_0)$; that is $\partial_1\partial_2 f(x_0, y_0) = \partial_2\partial_1 f(x_0, y_0)$. □

The concept of strong differentiation lead us to a satisfactory result, which we can summarize as follows:

(vi) If the strong derivative $\partial_2\partial_1 f(x_0, y_0)$ exists and it makes sense to consider the strong derivative $\partial_1\partial_2 f(x_0, y_0)$ (that is, $\partial_2 f$ exists in a set which accumulates at (x_0, y_0)), then the latter exists and they coincide.

2 Equality almost everywhere

We might also ask to what extent the equality of mixed partial derivatives holds for functions which admit those second derivatives almost everywhere. Some important results have been obtained for convex functions. A well known result by Alexandrov establishes that convex functions admit a generalized Peano derivative of order 2 in the sense that for almost every x

$$f(x+h) = f(x) + L(h) + A(h,h) + o_x(|h|^2),$$

where L is a linear map and A is a quadratic form. Less clear is whether A can be obtained from the differentiation of the generalized differential of f and whether such double differentiation gives a symmetric Hessian. The affirmative answer to this question has been established by Rockafellar [18].

The Russian mathematician G. P. Tolstov clarified several questions related to the equality of mixed derivatives in two papers published in 1949. Unfortunately, only one of those articles was translated into English [19], so that the interesting results contained in the other paper [20] have been largely overlooked by the mathematical community. Most space in those papers is devoted to the construction of counterexamples. In fact, he proved [20]:

4. There exists a function $f \in C^1(O,\mathbb{R})$, the mixed second derivatives of which exist at every point of O, but such that $\partial_2\partial_1 f \neq \partial_1\partial_2 f$ on a set $P \subset O$ of positive measure.

5. There exists a function $f \in C^1(O,\mathbb{R})$, the mixed second derivatives of which exist almost everywhere in O, and such that $\partial_2\partial_1 f \neq \partial_1\partial_2 f$ almost everywhere in O.

On the positive direction, he improved Young's theorem as follows [19]:

6. If the function f has all second derivatives everywhere in O, then the equality of mixed derivatives holds in O.

This theorem with *existence* replaced by *existence almost everywhere* in both the hypothesis and thesis had been already proved by Currier [21].

These type of results have still an undesirable feature, for they place conditions on the existence of the double derivatives $\partial_1^2 f$ and $\partial_2^2 f$. Actually, Tolstov obtained some results which do not place conditions on the homogeneous second derivatives. In the remainder of this work, we shall review and develop them. In particular, we shall stress the importance for applications of the Lipschitz conditions on the first derivatives.

We start with an important lemma from Tolstov's paper. Since there are no published English translations, we provide the proof.

Lemma 4 (Tolstov [20]). *Let $O = (a,b) \times (c,d) \subset \mathbb{R}^2$ and let*

$$f(x,y) = \int_a^x \mathrm{d}u \int_c^y h(u,v)\mathrm{d}v,$$

where $h \in L^1(\bar{O}, \mathbb{R})$. Then there is a measurable set $e_1 \subset (a, b)$ with $|e_1| = b - a$ such that for every $x \in e_1$ and for every $y \in (c, d)$

$$\partial_1 f(x, y) = \int_c^y h(x, v) \mathrm{d}v. \tag{2}$$

Clearly, by Fubini's theorem, the integrals in the definition of $f(x, y)$ can be exchanged, and hence a similar statement holds for the derivative with respect to y. Fubini's theorem will play a very important role in the proof of this lemma and in the proofs of the next theorems. The reader is referred to Aksoy and Martelli [22] for a discussion of the relationship between Fubini's and Schwarz's theorems.

PROOF. Differentiating $f(x, y)$ with respect to x, we obtain for $x \in X_y \subset [a, b]$ with $|X_y| = b - a$, (Fundamental Theorem of Calculus e.g. [23, Theor. 8.17])

$$\partial_1 f(x, y) = \int_c^y h(x, v)\, \mathrm{d}v. \tag{3}$$

Let $\Lambda = \cup_y (X_y \times \{y\})$, so that $|\Lambda| = (b - a)(d - c)$. Equation (3) holds for $(x, y) \in \Lambda$. Let $Y_x \subset (c, d)$ be the coordinate slices defined by $\{x\} \times Y_x = (\{x\} \times (c, d)) \cap \Lambda$, or equivalently, $Y_x = \pi_2(\pi_1^{-1}(x) \cap \Lambda)$. Fubini's theorem applied to the characteristic function of Λ gives that there is some $e_1 \subset (a, b)$, $|e_1| = b - a$, such that for every $x \in e_1$, $|Y_x| = c - d$. Let $x \in e_1$. For $y \in Y_x$ Eq. (3) is true. We wish to show that it holds for any $y \in (c, d)$. Let $y \in (c, d)$ and let h_n be any sequence converging to zero. Let

$$\varphi_n^{\pm}(y) := \frac{1}{h_n} \int_x^{x+h_n} \mathrm{d}u \int_c^y h^{\pm}(u, v) \mathrm{d}v.$$

where h^+ and h^- are the positive and negative parts of h, respectively. Since the functions $\varphi_n^+(y)$ are monotone and continuous and converge in a dense subset (for $n \to \infty$) to the continuous function $\int_c^y h^+(x, v)\, \mathrm{d}v$, they do the same everywhere on (c, d), and analogously for $\varphi_n^-(y)$. By the arbitrariness of h_n, Eq. (3) is true for every $y \in (c, d)$, provided $x \in e_1$. □

Remark 5. Notice that the Fundamental Theorem of Calculus (e.g. [23, Theor. 8.17]) states that Eq. (2) is true for $x \in e_1(y)$, where $|e_1(y)| = b - a$, but $e_1(y)$ might depend on y. The previous lemma states that e_1 does not depend on y.

A corollary is:

Theorem 6 (Tolstov [20])**.** *Let h, f and O be as in Lemma 4 above. There are measurable sets e_1 and e_2 with $|e_1| = b - a$, $|e_2| = d - c$, such that* *(a1) Anywhere in* $\{x\} \times (c,d)$ *with* $x \in e_1$*, we have*

$$\partial_1 f(x,y) = \int_c^y h(x,v)\mathrm{d}v. \tag{4}$$

(a2) Anywhere in $(a,b) \times \{y\}$ *with* $y \in e_2$*, we have*

$$\partial_2 f(x,y) = \int_a^x h(u,y)\mathrm{d}u. \tag{5}$$

(b) There exists a measurable set $E \subset e_1 \times e_2$ *with* $|E| = (b-a)(d-c)$*, such that for every* $(x,y) \in E$*, the mixed derivatives exist, and*

$$\partial_2\partial_1 f(x,y) = h(x,y) = \partial_1\partial_2 f(x,y). \tag{6}$$

Moreover, for every $x \in e_1$*,* $|\pi_1^{-1}(x) \cap E| = d-c$*, and for every* $y \in e_2$*,* $|\pi_2^{-1}(y) \cap E| = b-a$*.*

With respect to Tolstov's paper, we have included the last statement of point (b). Although this inclusion lengthens the proof, we give this version in order to be as complete as possible. A similar statement could be included at the end of the next theorems.

Proof. Let e_1 and e_2 be as in Lemma 4. Then (a1) and (a2) are a rephrasing of that lemma. Differentiating Eq. (4) with $x \in e_1$ with respect to y, we obtain that there is $e^2(x) \subset (c,d)$, $|e^2(x)| = d-c$ such that for $y \in e^2(x)$, the derivative $\partial_2\partial_1 f(x,y)$ exists and

$$\partial_2\partial_1 f(x,y) = h(x,y). \tag{7}$$

By taking the intersection of $e^2(x)$ with e_2, if necessary, we can assume that $e^2(x) \subset e_2$.

Let $E^1 = \cup_{x\in e_1}\{x\} \times e^2(x)$, so that $|E^1| = (b-a)(d-c)$ and on E^1 Eq. (7) holds true. Observe that $\pi_1(E^1) \subset e_1$ and $\pi_2(E^1) \subset e_2$. By Fubini's theorem, there is $c_2 \subset e_2$, $|c_2| = d-c$, such that for every $y \in c_2$, $d_1(y) := \pi_1(\pi_2^{-1}(y) \cap E^1) \subset e_1$, is such that $|d_1(y)| = b-a$.

By taking the intersection of $e^2(x)$ with c_2, if necessary, we can assume that $e^2(x) \subset c_2$. This redefinition does not change the properties of E^1, which gets replaced as follows: $E^1 \to E^1 \cap (e_1 \times c_2)$, but now $\pi_2(E^1) \subset c_2$, and for every $y \in c_2$, $d_1(y) = \pi_1(\pi_2^{-1}(y) \cap E^1) \subset e_1$ is such that $|d_1(y)| = b-a$.

Analogously, starting from Eq. (5) and working with the roles of x and y exchanged, we obtain that there is $c_1 \subset e_1$, $|c_1| = b-a$, such that for every

$y \in e_2$, there is $e^1(y) \subset c_1$, $|e^1(y)| = b-a$, such that on $E^2 = \cup_{y\in e_2} e^1(y) \times \{y\}$ the derivative $\partial_1\partial_2 f(x,y)$ exists and

$$\partial_1\partial_2 f(x,y) = h(x,y). \tag{8}$$

Moreover, $\pi_1(E^2) \subset c_1$ and for every $x \in c_1$, $d_2(x) := \pi_2(\pi_1^{-1}(x) \cap E^2) \subset e_2$ is such that $|d_2(x)| = d - c$.

Let us define $E = E^1 \cap E^2$. Then $E \subset c_1 \times c_2$, and for every $x \in c_1$,

$$\pi_2(\pi_1^{-1}(x) \cap E) = \pi_2(\pi_1^{-1}(x) \cap E^1) \cap \pi_2(\pi_1^{-1}(x) \cap E^2) = e^2(x) \cap d_2(x),$$

where both sets on the righthand side have full measure; thus

$$|\pi_2(\pi_1^{-1}(x) \cap E)| = d - c$$

(and analogously for the analogous statement with x and y exchanged). Finally, (7) and (8) are true on E, which proves (b), keeping (a1) and (a2) once we redefine $c_1 \to e_1$, $c_2 \to e_2$. □

We can also obtain a related theorem which adds information on the differentiability properties of f:

Theorem 7. *Let $f\colon [a,b] \times [c,d] \to \mathbb{R}$ be such that $f(x,\cdot) : [c,d] \to \mathbb{R}$ and $f(\cdot,y) : [a,b] \to \mathbb{R}$ are absolutely continuous for every $x \in [a,b]$ and $y \in [c,d]$, respectively. The following properties are equivalent:*

(i) There is $e_1 \subset (a,b)$, $|e_1| = b-a$, such that for $x \in e_1$, $\partial_1 f(x,\cdot)$ exists for every y. Moreover, it is absolutely continuous over $[c,d]$, and $\partial_2\partial_1 f \in L^1([a,b] \times [c,d])$.

(ii) There is $e_2 \subset (c,d)$, $|e_2| = d-c$, such that for $y \in e_2$, $\partial_2 f(\cdot,y)$ exists for every x. Moreover, it is absolutely continuous over $[a,b]$, and $\partial_1\partial_2 f \in L^1([a,b] \times [c,d])$.

Suppose they hold true. Then there is a subset $E \subset e_1 \times e_2$, $|E| = (b-a)(d-c)$, such that on E the function f is differentiable, $\partial_2\partial_1 f(x,y)$, $\partial_1\partial_2 f(x,y)$ exist, and

$$\partial_2\partial_1 f = \partial_1\partial_2 f.$$

Proof. Assume (i). For $x \in e_1$, the function $\partial_1 f(x,\cdot)$ is absolutely continuous. Thus, for every y,

$$\partial_1 f(x,y) - \partial_1 f(x,c) = \int_c^y \partial_2\partial_1 f(x,v)\mathrm{d}v. \tag{9}$$

Since $f(\cdot, y)$ is absolutely continuous, we obtain upon integration

$$f(x,y) - f(a,y) - f(x,c) + f(a,c) = \int_a^x \mathrm{d}u \int_c^y \partial_2\partial_1 f(u,v)\mathrm{d}v.$$

By Theorem 6 applied to the righthand side, there is a subset $\tilde{e}_1 \subset (a,b)$, $|\tilde{e}_1| = b-a$, such that Eq. (9) holds. This is already known to be true with $\tilde{e}_1 = e_1$. The same theorem establishes the existence of $e_2 \subset (c,d)$, $|e_2| = d-c$, such that for $y \in e_2$ and for every $x \in (a,b)$,

$$\partial_2 f(x,y) - \partial_2 f(a,y) = \int_a^x \partial_2\partial_1 f(u,y)\mathrm{d}u.$$

This last equation shows that for $y \in e_2$, the function $\partial_2 f(\cdot, y)$ is absolutely continuous. Thus for every $y \in e_2$, $|e_2| = d-c$, there is $e_1(y) \subset (a,b)$, $|e_1(y)| = b-a$, such that for $x \in e_1(y)$, and hence for almost every pair $(x,y) \in (a,b)\times(c,d)$, we have $\partial_1\partial_2 f(x,y) = \partial_2\partial_1 f(x,y)$. This implies that $\partial_1\partial_2 f \in L^1([a,b]\times[c,d])$; that is, (ii) is true. The proof that (ii) implies (i) is analogous.

It remains only to prove the differentiability of f on E, the remaining part of the last statement being an immediate consequence of Theorem 6.

Let us prove the differentiability of f at $(x_0, y_0) \in E$. The partial derivative $\partial_2 f(x, y_0)$ exists for every x and is absolutely continuous in x. Analogously, $\partial_1 f(x_0, y)$ exists for every y and is absolutely continuous in y. We have

$$\begin{aligned}
&f(x_0+\Delta x, y_0+\Delta y) - f(x_0,y_0)\\
&= [f(x_0+\Delta x, y_0+\Delta y) - f(x_0+\Delta x, y_0)] + [f(x_0+\Delta x, y_0) - f(x_0,y_0)]\\
&= \partial_2 f(x_0+\Delta x, y_0)\Delta y + o_2(\Delta y) + \partial_1 f(x_0,y_0)\Delta x + o_1(\Delta x)\\
&= [\partial_2 f(x_0,y_0) + \partial_1\partial_2 f(x_0,y_0)\Delta x + o_3(\Delta x)]\Delta y + o_2(\Delta y)\\
&\quad + \partial_1 f(x_0,y_0)\Delta x + o_1(\Delta x)\\
&= \partial_2 f(x_0,y_0)\Delta y + \partial_1 f(x_0,y_0)\Delta x + R(\Delta x, \Delta y),
\end{aligned}$$

where $R(\Delta x, \Delta y)/(\Delta x^2 + \Delta y^2)^{1/2} \to 0$ for the denominator going to zero. □

Remark 8. It is well known that in the theory of distributions, the equality of mixed derivatives holds at any order of differentiation [24]. In order to convert this fact into a claim for ordinary differentiation it is necessary that the second derivatives $\partial_2\partial_1 f$ and $\partial_1\partial_2 f$ be *regular* distributions, namely representable as the integral of the test function φ with $L^1(\bar{O}, \mathbb{R})$ functions. Tolstov's Lemma allows us to remove this double condition on the second mixed derivatives, for it is sufficient to place that condition on just $\partial_2\partial_1 f$.

3 Lipschitz conditions on the partial derivatives

Let us recall that a function $g\colon U \to \mathbb{R}^k$ defined on an open set $U \subset \mathbb{R}^n$ is Lipschitz if for every $p, q \in U$,

$$\|g(p) - g(q)\| < K\|p - q\|$$

for some $K > 0$. It is locally Lipschitz if this inequality holds over every compact subset of U, with K dependent on the compact subset.

A function $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^2$, $(x, y) \mapsto f(x, y)$, is differentiable with Lipschitz differential, or $C^{1,1}$ for short, if $df\colon U \to \mathbb{R}^2$ is Lipschitz. Clearly, if $f \in C^{1,1}$ with Lipschitz constant K, then the partial derivative $\partial_1 f(x, \cdot)$ regarded as a function of y is K-Lipschitz. In particular, the Lipschitz constant does not change if we change x. In this case, we say that $\partial_1 f(x, y)$ is Lipschitz in y uniformly in x. Analogously, $\partial_2 f(x, y)$ is Lipschitz in x uniformly in y, and two other similar combinations hold.

For functions admitting Lipschitz partial derivatives, the L^1 condition on the mixed derivative which we met in Theorem 7 is satisfied.

Theorem 9. *Let $f\colon [a, b] \times [c, d] \to \mathbb{R}$, be such that $f(x, \cdot) : [c, d] \to \mathbb{R}$ and $f(\cdot, y) : [a, b] \to \mathbb{R}$ are absolutely continuous for every $x \in [a, b]$ and $y \in [c, d]$, respectively. The following properties are equivalent:*

(i) There is $e_1 \subset (a, b)$, $|e_1| = b - a$, such that for $x \in e_1$, $\partial_1 f(x, \cdot)$ exists for every y, and moreover, it is Lipschitz over $[c, d]$ uniformly for $x \in e_1$.

(ii) There is $e_2 \subset (c, d)$, $|e_2| = d - c$, such that for $y \in e_2$, $\partial_2 f(\cdot, y)$ exists for every x, and moreover, it is Lipschitz over $[a, b]$ uniformly for $y \in e_2$.

Suppose they hold true. Then there is a subset $E \subset e_1 \times e_2$, $|E| = (b - a)(d - c)$, such that on E the function f is differentiable, $\partial_2\partial_1 f(x, y)$, $\partial_1\partial_2 f(x, y)$ exist and are bounded and

$$\partial_2\partial_1 f = \partial_1\partial_2 f.$$

PROOF. Suppose (i) is true and let $x \in e_1$ so that $\partial_1 f(x, \cdot)$ exists and is K-Lipschitz over $[c, d]$. Then $|\partial_2\partial_1 f(x, \cdot)| \le K$ a.e. in (c, d), and by the assumption of Lipschitz uniformity, this bound holds for every $x \in e_1$. In particular, $|\partial_2\partial_1 f| \le K$ holds almost everywhere on O. Thus $\partial_2\partial_1 f \in L^\infty(\bar{O}, \mathbb{R}) \subset L^1(\bar{O}, \mathbb{R})$ and condition (i) of Theorem 7 is satisfied. In particular, the last statement of that theorem implies that "$|\partial_2\partial_1 f| \le K$, f is differentiable and $\partial_2\partial_1 f$, $\partial_1\partial_2 f$ exist and coincide" almost everywhere on a subset $W \subset O$, $|W| = (b - a)(d - c)$.

By Theorem 7, there is also a set e_2 such that if $y \in e_2$, then $\partial_2 f(\cdot, y)$ exists for every x, and moreover, it is absolutely continuous; thus

$$\partial_2 f(x,y) - \partial_2 f(a,y) = \int_a^x \partial_1 \partial_2 f(u,y) \mathrm{d}u.$$

However, by Fubini's theorem, there is $b_2 \subset (c,d)$, $|b_2| = d-c$, such that for every $y \in b_2$, $|\pi^{-1}(y) \cap W| = b-a$. Thus for almost every y, namely for $y \in e_2 \cap b_2$, we have that $\partial_2 f(\cdot, y)$ exists for every x and it is absolutely continuous, and for almost every x, we have "$\partial_1\partial_2 f = \partial_2\partial_1 f$ and $|\partial_2\partial_1 f| \le K$." Thus for $y \in e_2 \cap b_2$, $\partial_2 f(\cdot, y)$ is K Lipschitz over (a,b), where K does not depend on y. Thus (ii) is proved once we rename $e_2 \cap b_2$ as e_2.

The last statement follows from the previous paragraph or from the last one in Theorem 7. □

For C^1 functions we have:

Theorem 10. *Let Ω be an open subset of $\mathbb{R}^2$ and let $f \in C^1(\Omega, \mathbb{R})$. Then the following conditions are equivalent:*

(i) For every x, the partial derivative $\partial_1 f(x, \cdot)$ is locally Lipschitz, locally uniformly with respect to x.

(ii) For every y, the partial derivative $\partial_2 f(\cdot, y)$ is locally Lipschitz, locally uniformly with respect to y.

If they hold true, for instance $f \in C^{1,1}_{loc}(\Omega, \mathbb{R})$, then on a set $E \subset \Omega$, $|\Omega \backslash E| = 0$, $\partial_2\partial_1 f$ and $\partial_1\partial_2 f$ exist, f is differentiable and $\partial_2\partial_1 f = \partial_1\partial_2 f$. In particular, $\partial_2\partial_1 f$ and $\partial_1\partial_2 f$ belong to $L^\infty_{loc}(\Omega, \mathbb{R})$.

Proof. Let $p \in \Omega$ and let us consider an open neighborhood $(a,b) \times (c,d)$ of p such that $\bar{O} \subset \Omega$. Let us assume (i). By Theorem 9, we need only to show that $\partial_2 f(\cdot, y)$ is K-Lipschitz in (a,b) for every chosen value of $y \in (c,d)$, provided it is so for almost every $y \in (c,d)$. Let $y \in (c,d)$ and let $\epsilon > 0$. The function $\partial_2 f$ is continuous, thus uniformly continuous over the compact set $[a,b] \times [c,d]$. We can find a $\delta > 0$ such that whenever $|y_2 - y_1| < \delta$ with $y_1, y_2 \in [c,d]$, we have $|\partial_2 f(x, y_2) - \partial_2 f(x, y_1)| < \epsilon$ for every $x \in [a,b]$. We can find some $\bar{y} \in (y-\delta, y+\delta) \cap [c,d]$ such that $\partial_2 f(\cdot, \bar{y})$ is K-Lipschitz. Thus

$$\begin{aligned}|\partial_2 f(x_2, y) - \partial_2 f(x_1, y)| &\le |\partial_2 f(x_2, \bar{y}) - \partial_2 f(x_1, \bar{y})| + |\partial_2 f(x_2, \bar{y}) - \partial_2 f(x_2, y)| \\ &\quad + |\partial_2 f(x_1, \bar{y}) - \partial_2 f(x_1, y)| \le K\|x_2 - x_1\| + 2\epsilon.\end{aligned}$$

From the arbitrariness of ϵ, x_1 and x_2, we obtain that $\partial_2 f(\cdot, y)$ is K-Lipschitz for every chosen value of y. The remaining claims follow trivially from Theorem 9. □

We stress that if $f \in C^1$, then the fact that $\partial_1 f(x,y)$ is Lipschitz in y uniformly in x, and that $\partial_2 f(x,y)$ is Lipschitz in x uniformly in y, does not guarantee that $f \in C^{1,1}$; it is sufficient to consider the function $f(x,y) = |x|^{3/2}$. As a consequence, the assumptions of this theorem are weaker than $f \in C^{1,1}_{loc}(\Omega, \mathbb{R})$.

For the $C^{1,1}_{loc}(\Omega, \mathbb{R})$ case, the equality of mixed derivatives can also be obtained as a consequence of Young's (point 3 above) and Rademacher's theorems. We recall that the latter states that every Lipschitz function is almost everywhere differentiable [25]. Indeed:

Theorem 11. *Let $f\colon \Omega \to \mathbb{R}$, $f \in C^{1,1}_{loc}$. Then f is twice differentiable almost everywhere and in such differentiability set $\partial_2\partial_1 f = \partial_1\partial_2 f$.*

PROOF. The differential $df\colon \Omega \to \mathbb{R}^2$ is Lipschitz, thus differentiable almost everywhere (Rademacher's theorem). If p belongs to the differentiability set, then $\partial_1 f$ and $\partial_2 f$, being components of the differential, are there differentiable. Thus by Young's theorem 3., we have $\partial_2\partial_1 f = \partial_1\partial_2 f$ at p. □

4 Some applications

In this section we explore some applications that motivated our study. They are in the area of differential geometry, but it is likely that many other applications can be found.

4.1 Usefulness of Lipschitz one-forms

A rather natural application of these results is in the study of Lipschitz 1-forms, namely 1-forms with Lipschitz coefficients, over differentiable manifolds (at least $C^{1,1}$). Indeed, some related results have been already developed following paths independent of the above considerations.

If $f \in C^{1,1}_{loc}$, then by Theorem 11, $\mathrm{d}^2 f = 0$ almost everywhere in the Lebesgue 2-dimensional measure of any 2-dimensional $C^{1,1}$ embedded manifold. Thus the exterior differential satisfies $\mathrm{d}^2 = 0$ in a well defined sense whenever 0-forms and 1-forms are taken with the correct degree of differentiability. In particular, if ω is a Lipschitz 1-form, then Stokes theorem

$$\int_S \mathrm{d}\omega = \int_{\partial S} \omega$$

still holds true [26].

One also expects that Lipschitz distributions of hyperplanes should be integrable according to the usual rule for C^1 distributions. Namely, let ω be a Lipschitz 1-form. Then the distribution $\mathrm{Ker}\,\omega$ should be integrable if and only if $\omega \wedge \mathrm{d}\omega = 0$. This result has indeed been proved [26, 27].

The nice behavior of locally Lipschitz 1-forms suggests to study (pseudo-) Riemannian $C^{2,1}$ manifolds endowed with Lipschitz connections ∇ and $C^{1,1}$ metrics. Indeed, in Cartan's approach, the connection is regarded as a Lie algebra-valued 1-form in the bundle of reference frames. In such framework, the Riemannian tensor would be a locally bounded element of $L^{\infty}_{loc}(M)$ and hence would be defined only almost everywhere in the Lebesgue 2-dimensional measure. In particular, it could be discontinuous though locally bounded. This feature would be quite appreciated in the theory of Einstein's general relativity. There the Ricci tensor is proportional to the stress-energy tensor, and so it is necessarily discontinuous for the typical mass distribution of a planet; the reader may consider the discontinuity in density which takes place at the planet's boundary.

4.2 An application to mathematical relativity

As mentioned, the assumptions of Theorem 10 are weaker than the condition $f \in C^{1,1}_{loc}$. We wish to describe shortly an example of application where those weaker conditions turn out to be important.

In Einstein's General Relativity, the spacetime continuum is represented with a Lorentzian manifold [28], namely a differentiable manifold endowed with a metric of signature $(-,+,+,+)$. Observers or massive particles are represented by C^1 curves $x(s)$ which are timelike, namely such that $g(x',x') < 0$. Unfortunately, for various mathematical arguments it is necessary to consider limits of such curves, and those limits are rarely C^1, but are necessarily Lipschitz. Lipschitz mathematical objects arise quite naturally in General Relativity, ultimately because the light cones on spacetime (at $p \in M$ the light cone is given by the subset of T_pM where g vanishes) place a bound on the local speed of massive objects.

A typical problem which is met in the discussion of the clock effect (or twin paradox) deals with curve variations $x(t,s)$ of timelike geodesics $x(\cdot,s)$ parametrized by s, where the transverse curves $x(t,\cdot)$ are just Lipschitz. In this situation, the properties of the exponential map allow one to prove that the tangent $\partial_t x(t,s)$ is Lipschitz in s uniformly in t, exactly the assumptions of Theorem 10 (the function need not be differentiable in s). It turns out that in order to prove formulas such as the first variation formula for the energy functional of differential geometry,

$$E[x] = \frac{1}{2}\int_0^1 g(\dot{x},\dot{x})\mathrm{d}t,$$

one needs to switch $\partial_s \partial_t x$ for $\partial_t \partial_s x$, an operation which is indeed allowed thanks to Theorem 10. Thus, this theorem can be used to operate with Lipschitz curves in Lorentzian (or Riemannian) geometry much in the same way as it is usually done with C^1 curves [10].

5 Conclusions

We have reviewed the notion of strong differentiation and Mikusiński's result on the equality of mixed partial derivatives. The assumptions do not demand the existence and continuity of any second derivative in a neighborhood of the point. Rather, the theorem assumes the weaker notion of strong differentiability of one first derivative at the point. This possibility was suggested by previous applications of the concept of strong differentiation where it proved to be particularly advantageous; e.g. the inverse function theorem.

We then considered results which prove the existence and equality of mixed partial derivatives almost everywhere. We have presented and elaborated previous results by Tolstov stressing the importance of the Lipschitz condition on first partial derivatives for applications. The advantage of this approach over alternative distributional approaches becomes clear whenever one cannot conclude that both mixed second partial derivatives are summable. We have ended this work giving a specific example of application where this classical approach is more effective and justified.

Acknowledgment. I thank an anonymous referee for some useful criticisms. Work partially supported by GNFM of INDAM.

References

[1] T. J. Higgins, *A note on the history of mixed partial derivatives*, Scripta Mathematica, **7** (1940), 59–62.

[2] H. A. Schwarz, *Communication*, Archives des Sciences Physiques et Naturelles, **48** (1873), 38–44.

[3] U. Dini, *Lezioni di analisi infinitesimale, Vol. I*, Stab. Tipogradico Succ. Fratelli Nistri, Pisa, 1907.

[4] G. Peano, *Sur l'interversion des dérivations partielles*, Mathesis, **10** (1890), 153–154.

[5] W. Rudin, *Principles of mathematical analysis*, McGraw-Hill, New York, 1976.

[6] N. Kimura, J. Burr, *Discontinuous function with partial derivatives everywhere*, Amer. Math. Monthly, **67** (1960), 813–814.

[7] T. M. Apostol, *Mathematical Analysis*, Addison-Wesley Publishing Company, Reading, 1974.

[8] G. Peano, *Sur la définition de la dérivée*, Mathesis, **2** (1892), 12–14.

[9] S. Dolecki, G. H. Greco, *Amazing oblivion of Peano's mathematical legacy* (2012).

[10] E. Minguzzi, *Convex neighborhoods for Lipschitz connections and sprays* (2013), ArXiv:1308.6675.

[11] M. Esser, O. Shisha, *A modified differentiation*, Amer. Math. Monthly, **71** (1964), 904–906.

[12] A. Nijenhuis, *Strong derivatives and inverse mappings*, Amer. Math. Monthly, **81** (1974), 969–980.

[13] E. B. Leach, *A note on inverse function theorems*, Proc. Amer. Math. Soc., **12** (1961), 694–697.

[14] J. Mikusiński, *On partial derivatives*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astr. Phys., **20** (1972), 941–944.

[15] J. Mikusiński, *O pochidnych ezqstkowych i pelnych* (*On partial and full derivatives*,in Polish), Roczniki Polskiego Towarzystwa Matematycznego. Seria II: Wiadomosci Matematyczne, **16** (1973), 23–35.

[16] J. Mikusiński, *The Bochner integral*, Academic Press, New York, 1978.

[17] K. Skórnik, *On full derivatives*, Ann. Polon. Math., **43** (1983), 323–325.

[18] R. T. Rockafellar, *Second-order convex analysis*, J. Nonlinear Convex Anal., **1** (1999), 1–16.

[19] G. P. Tolstov, *On partial derivatives* (Russian), Izvestiya Akad. Nauk SSSR. Ser. Mat., **13** (1949), 425–426. Reprinted: Tolstov, *On partial derivatives*, Translations, Ser. 1, Vol. 10: Functional Analysis and Measure Theory, Amer. Math. Soc., 1962.

[20] G. P. Tolstov, *On the mixed second derivative* (Russian), Mat. Sbornik (N.S.), **24(66)** (1949), 27–51.

[21] A. E. Currier, *Proof of the fundamental theorems on second-order cross partial derivatives*, Trans. Amer. Math. Soc., **35** (1933), 245–253.

[22] A. Aksoy, M. Martelli, *Mixed partial derivatives and Fubini's theorem*, College Math. J., **33** (2002), 126–130.

[23] W. Rudin, *Real and complex analysis*, McGraw-Hill, London, 1970.

[24] V. S. Vladimirov, *Methods of the theory of generalized functions*, Taylor and Francis, London, 2002.

[25] L. C. Evans, R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, 1992.

[26] S. Simić, *Lipschitz distribution and Anosov flows*, Proc. Amer. Math. Soc., **124** (1996), 1869–1877.

[27] F. Rampazzo, *Frobenius-type theorems for Lipschitz distributions*, J. Differential Equations, **243** (2007), 270–300.

[28] S. W. Hawking, G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Univ. Press, Cambridge, 1973.