Rev. Mat. Iberoam. **30** (2014), no. 1, 349–367
DOI 10.4171/RMI/782 

# On the $L^p$-differentiability of certain classes of functions

Giovanni Alberti, Stefano Bianchini and Gianluca Crippa

**Abstract.** We prove the $L^p$-differentiability at almost every point for convolution products on $\mathbb{R}^d$ of the form $K*\mu$, where $\mu$ is bounded measure and $K$ is a homogeneous kernel of degree $1-d$. From this result we derive the $L^p$-differentiability for vector fields on $\mathbb{R}^d$ whose curl and divergence are measures, and also for vector fields with bounded deformation.

## 1. Introduction

Let $u$ be a convolution product on $\mathbb{R}^d$ of the form

$$u := K * \mu\,, \tag{1.1}$$

where $\mu$ is a bounded measure[1] and $K$ is a kernel of class $C^2$ away from 0 and homogeneous of degree $1-d$. The main result of this paper (Theorem 3.4) states that $u$ is differentiable in the $L^p$ sense[2] at almost every point for every $p$ with $1 \le p < \gamma(1)$, where $\gamma(q) := qd/(d-q)$ is the exponent of the Sobolev embedding for $W^{1,q}$ in dimension $d$.

Using this result, we show that a vector field $v$ on $\mathbb{R}^d$ is $L^p$-differentiable almost everywhere for the same range of $p$ if either of the following conditions holds (see Propositions 4.2 and 4.3):

(a) the (distributional) curl and divergence of $v$ are measures;

(b) $v$ belongs to the class $BD$ of maps with bounded deformation, that is, the (distributional) *symmetric* derivative $\frac{1}{2}(\nabla v + \nabla^t v)$ is a measure.

**Relation with Sobolev and $BV$ functions.** If the measures in the statements above are replaced by functions in $L^q$ for some $q > 1$, then $u$ and $v$ would be

*Mathematics Subject Classification* (2010): Primary 26B05; Secondary 42B20, 46E35.
*Keywords:* Approximate differentiability, Lusin property, convolution operators, singular integrals, Calderón–Zygmund decomposition, Sobolev functions, functions with bounded variation, functions with bounded deformation.

[1] That is, a measure with finite total mass.

[2] The definition of $L^p$-differentiability is recalled in §2.2.

(locally) in the Sobolev class $W^{1,q}$ (see Lemma 3.9), and it is well known that a function in this class is $L^{\gamma(q)}$-differentiable almost everywhere when $q < d$, and differentiable almost everywhere in the classical sense when $q > d$, see for instance Sections 6.1.2 and 6.2 in [6].[3]

Functions in the $BV$ class – namely those functions whose distributional derivative is a measure – share the same differentiability property of functions in the class $W^{1,1}$ (see Section 6.1.1 in [6]). Note, however, that the functions $u$ and $v$ that we considered above in general fail to be of class $BV$, even locally.[4]

**A Lusin-type theorem.** Consider a Lipschitz function $w$ on $\mathbb{R}^d$ whose (distributional) Laplacian is a measure. Thus $\nabla w$ satisfies assumption (a) above, and therefore is $L^1$-differentiable almost everywhere. Using this fact we can show that $w$ admits an $L^1$-Taylor expansion of order two at almost every point and consequently has the Lusin property with functions of class $C^2$ (see §2.4 and Proposition 4.4). This Lusin property is used in [1] to prove that $w$ has the so-called weak Sard property, and was the original motivation for this paper.

**Comparison with existing results.** The proof of Theorem 3.4 is based on classical arguments from the theory of singular integrals, but, somewhat surprisingly, we could not find this statement in the literature.

There are, however, a few results which are closely related: the approximate differentiability [5] at almost every point of the convolution product in (1.1) was already proved in Theorem 6 of [9].[6] It should be noted that the notion of approximate differentiability is substantially weaker than $L^1$-differentiability; in particular, in Remark 4.7 we show that the result in [9] cannot be used to prove the Lusin property mentioned in the previous paragraph.

The $L^1$-differentiability of $BD$ functions was first proved in [2] (see Theorem 7.4). This proof is quite different from ours and, as far as we can see, cannot be adapted to the more general setting considered in Theorem 3.4.

**Optimality of the exponent $p$.** The range of $p$ for which we can prove $L^p$-differentiability is optimal in all cases considered above, with the exception of $BD$ functions (see Remarks 3.5 and 4.5).

The paper is organized as follows: in Section 2 we introduce the notation and recall a few basic facts on differentiability in the $L^p$ sense, in Section 3 we state and prove the main result (Theorem 3.4), and in Section 4 we derive a few applications.

**Acknowledgment.** We thank Piotr Hajłasz for many helpful comments.

---

[3] For $q > d$ the result refers to the continuous representative of the function.

[4] An example of $u := K * \mu$ which is not (locally) $BV$ is obtained by taking $K(x) := |x|^{1-d}$ and $\mu$ equal to the Dirac mass at 0. An example of vector field with vanishing curl and measure divergence which is not (locally) $BV$ is the derivative of the fundamental solutions of the Laplacian, see §4.1. The existence of vector fields which are in $BD$ but not in $BV$ is less immediate, and is derived by the failure of Korn inequality for the exponent $p = 1$ proved in [10] (see also Section 2 in [5]).

[5] The definition of approximate differentiability is recalled in Remark 2.3 (v).

[6] For the special case $K(x) := |x|^{1-d}$ and $\mu$ replaced by a function in $L^1$ a sketch of proof was also given in a remark at page 129 of [3].

## 2. Notation and preliminary results

**2.1. Notation.** For the rest of this paper, $d \geq 2$ is a fixed integer. Sets and functions are tacitly assumed to be Borel measurable, and measures are always defined on the Borel $\sigma$-algebra. We use the following notation:

| | |
|---|---|
| $\mathrm{diam}(E)$ | diameter of the set $E$; |
| $\mathrm{conv}(E)$ | convex hull of the set $E$; |
| $\mathrm{dist}(E_1, E_2)$ | distance between the sets $E_1$ and $E_2$, that is, the infimum of $\lvert x_1 - x_2\rvert$ among all $x_1 \in E_1$, $x_2 \in E_2$; |
| $1_E$ | characteristic function of the set $E$ (valued in $\{0,1\}$); |
| $B(x,\rho)$ | open ball in $\mathbb{R}^d$ with radius $\rho$ and center $x$; |
| $B(\rho)$ | open ball in $\mathbb{R}^d$ with radius $\rho$ and center 0; |
| $S^{d-1}$ | $:= \{x \in \mathbb{R}^d : \lvert x\rvert = 1\}$, unit sphere in $\mathbb{R}^d$; |
| $⨍_E f\, d\mu$ | $:= \frac{1}{\mu(E)} \int_E f\, d\mu$, average of the function $f$ over the set $E$ with respect to the positive measure $\mu$; |
| $\rho \cdot \mu$ | measure associated to the measure $\mu$ and the density function $\rho$, that is, $[\rho \cdot \mu](E) := \int_E \rho\, d\mu$ for every Borel set $E$; |
| $1_E \cdot \mu$ | restriction of the measure $\mu$ to the set $E$; |
| $\lvert\mu\rvert$ | positive measure associated to a real- or vector-valued measure $\mu$ (total variation); |
| $\lVert\mu\rVert$ | $:= \lvert\mu\rvert(\mathbb{R}^d)$, total mass of the measure $\mu$; |
| $\mathscr{L}^d$ | Lebesgue measure on $\mathbb{R}^d$; |
| $\mathscr{H}^k$ | $k$-dimensional Hausdorff measure (on any metric space); |
| $\omega_d$ | $:= \mathscr{L}^d(B(1))$, Lebesgue measure of the unit ball in $\mathbb{R}^d$; |
| $\gamma(q)$ | $:= qd/(d-q)$ for $1 \leq q < d$ and $\gamma(q) := +\infty$ for $q \geq d$; exponent of the Sobolev embedding for $W^{1,q}$ in dimension $d$. |

When the measure is not specified, it is assumed to be the Lebesgue measure, and in particular we often write $\int f(x)\, dx$ for the integral of $f$ with respect to $\mathscr{L}^d$.

As usual, we denote by $o(\rho^k)$ any real- or vector-valued function $g$ on $(0,+\infty)$ such that $\rho^{-k} g(\rho)$ tends to 0 as $\rho \to 0$, while $O(\rho^k)$ denotes any $g$ such that $\rho^{-k} g(\rho)$ is bounded in a neighbourhood of 0.

**2.2. Taylor expansions in the $L^p$ sense.** Let be $u$ a real function on $\mathbb{R}^d$. Given a point $x \in \mathbb{R}^d$, a real number $p \in [1,\infty)$, and an integer $k \geq 0$, we say that $u$ has a Taylor expansion of order $k$ in the $L^p$ sense at $x$, and we write $u \in t^{k,p}(x)$, if it can be decomposed as

$$u(x+h) = P^k_x(h) + R^k_x(h) \quad \text{for every } h \in \mathbb{R}^d, \tag{2.1}$$

where $P_x^k$ is a polynomial on $\mathbb{R}^d$ with degree at most $k$ and the remainder $R_x^k$ satisfies

$$\left[\fint_{B(\rho)} |R_x^k(h)|^p \, dh\right]^{1/p} = o(\rho^k)\,. \tag{2.2}$$

As usual, the polynomial $P_x^k$ is uniquely determined by (2.1) and the decay estimate (2.2).

When $u$ belongs to $t^{0,p}(x)$ we say that it has $L^p$-limit at $x$ equal to $P_x^0(0)$. When $u$ belongs to $t^{1,p}(x)$ we say that $u$ is $L^p$-differentiable at $x$ with derivative equal to the derivative of the polynomial $P_x^1$ at 0.

We write $u \in T^{k+1,p}(x)$ if the term $o(\rho^k)$ in (2.2) can be replaced by $O(\rho^{k+1})$. Accordingly, we write $u \in T^{0,p}(x)$ if

$$\left[\fint_{B(\rho)} |u(x+h)|^p \, dh\right]^{1/p} = O(1)\,.$$

The definitions above are given for real-valued functions defined on $\mathbb{R}^d$, but are extended with the necessary modifications to vector-valued functions defined on some open neighbourhood of the point $x$.

Finally, it is convenient to define $t^{k,\infty}(x)$ and $T^{k,\infty}(x)$ by replacing the left-hand side of (2.2) with the $L^\infty$ norm of $R_x^k(h)$ on $B(\rho)$. Note that $u$ belongs to $t^{k,\infty}(x)$ if and only if it agrees almost everywhere with a function which admits a Taylor expansion of order $k$ at $x$ in the classical sense.

**Remark 2.3.** (i) The space $t^{k,p}(x)$ and $T^{k,p}(x)$ were introduced in a slightly different form in [4] (see also Section 3.5 in [13]). The original definition differs from ours in that it also requires that the left-hand side of (2.2) is smaller that $c\rho^k$ for some finite constant $c$ and for *every* $\rho > 0$ (and not just for small $\rho$).[7]

(ii) The function spaces $t^{k,p}(x)$ and $T^{k,p}(x)$ satisfy (quite obviously) the inclusions $T^{k,p}(x) \subset T^{k,q}(x)$ and $t^{k,p}(x) \subset t^{k,q}(x)$ whenever $p \geq q$, and $T^{k+1,p}(x) \subset t^{k,p}(x) \subset T^{k,p}(x)$.

(iii) Concerning the last inclusion ($t^{k,p}(x) \subset T^{k,p}(x)$), the following nontrivial converse holds: if $u$ belongs to $T^{k,p}(x)$ for every $x$ in a set $E$, then $u$ belongs to $t^{k,p}(x)$ for *almost every* $x \in E$ (see Theorem 3.8.1 in [13]).[8]

(iv) We recall that function $u$ on $\mathbb{R}^d$ has approximate limit $a \in \mathbb{R}$ at $x$ if the set $\{h : |u(x+h) - a| \leq \varepsilon\}$ has density 1 at the point 0 for every $\varepsilon > 0$. It is immediate to check that if $u$ has $L^p$-limit equal to $a$ at $x$ for some $p \geq 1$, then it has also approximate limit $a$ at $x$.

(v) A function $u$ on $\mathbb{R}^d$ has approximate limit $a \in \mathbb{R}$ and approximate derivative $b \in \mathbb{R}^d$ at $x$ if the ratio $(u(x+h) - a - b \cdot h)/|h|$ has approximate limit 0 as $h \to 0$.

[7] This additional requirement is met if (and only if) the function $u$ satisfies the growth condition $\int_{B(\rho)} |u|^p \leq c\rho^{d+kp}$ for some finite $c$ and for sufficiently large $\rho$.

[8] For $k = 1$ this statement can be viewed as an $L^p$ version of the classical Rademacher theorem on the differentiability of Lipschitz functions. The fact that our definition of $t^{k,p}(x)$ and $T^{k,p}(x)$ differs from that considered in [13] has no consequences for the validity of this statement, which is purely local.

It is easy to check that if $u$ has $L^p$-derivative $b$ at $x$ then it also has approximate derivative $b$ at $x$.

**2.4. Lusin property.** Let $E$ be a set in $\mathbb{R}^d$ and $u$ a function defined at every point of $E$. We say that $u$ has the Lusin property with functions of class $C^k$ (on $E$) if for every $\varepsilon > 0$ there exists a function $v$ of class $C^k$ on $\mathbb{R}^d$ which agrees with $u$ in every point of $E$ except a subset with measure at most $\varepsilon$.

Note that $u$ has the Lusin property with functions of class $C^k$ provided that $u \in t^{k,1}(x)$ for a.e. $x \in E$, or, equivalently, $u \in T^{k,1}(x)$ for a.e. $x \in E$ (recall Remark 2.3 (iii)). Assume indeed that $E$ has finite measure: then for every $\varepsilon > 0$ we can find a compact subset $D$ such that $\mathscr{L}^d(E \setminus D) \leq \varepsilon$, $u$ is continuous on $D$, and estimate (2.2) holds *uniformly* for all $x \in D$,[9] and therefore the $L^p$-version of the Whitney extension theorem (see Theorem 3.6.3 in [13]) yields that $u$ agrees on $D$ with a function class $C^k$ on $\mathbb{R}^d$.

## 3. Differentiability of convolution products

**3.1. Assumptions on the convolution kernel.** Through the rest of this paper $K$ is a real function of class $C^2$ on $\mathbb{R}^d \setminus \{0\}$, homogeneous of degree $1-d$, that is, $K(\lambda x) = \lambda^{1-d} K(x)$ for every $x \neq 0$ and $\lambda > 0$.

It follows immediately that the derivative $\nabla K : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}^d$ is of class $C^1$ and homogeneous of degree $-d$. Moreover it satisfies the cancellation property

$$\int_{S^{d-1}} \nabla K \, d\mathscr{H}^{d-1} = 0 \,. \tag{3.1}$$

Indeed, let $a$ be the integral of $\nabla K$ over $S^{d-1}$, $\Omega$ the set of all $x \in \mathbb{R}^d$ such that $1 < |x| < 2$, $\nu$ the outer normal of $\partial\Omega$, and $e$ an arbitrary vector in $\mathbb{R}^d$. By applying the divergence theorem to the vector field $Ke$ and the domain $\Omega$, we obtain

$$\int_{\partial\Omega} K \, e \cdot \nu \, d\mathscr{H}^{d-1} = \int_{\Omega} \frac{\partial K}{\partial e} \, d\mathscr{L}^d \,.$$

Now, using the fact that $K$ is homogeneous of degree $1-d$ we obtain that the integral at the left-hand side is 0, while a simple computation shows that the integral at the right-hand side is equal to $a \cdot e$ times $\log 2$. Hence $a \cdot e = 0$, and since $e$ is arbitrary, $a = 0$.

**3.2. A first convolution operator.** Take $K$ as in the previous paragraph, and let $\mu$ be a bounded real-valued measure on $\mathbb{R}^d$. We define the convolution product $K * \mu$ by the usual formula

$$K * \mu(x) := \int_{\mathbb{R}^d} K(x-y) \, d\mu(y) \,. \tag{3.2}$$

[9] That is, the functions $g_\rho(x) := \rho^{-k} \int_{B(\rho)} |R^k_x(h)| \, dh$ converge uniformly to 0 as $\rho \to 0$. The existence of such $D$ is an easy consequence of Lusin theorem (for the continuity of $u$ on $D$) and Egorov theorem (for the uniform convergence of $g_\rho$ on $D$).

Since $|K(x)| \leq c|x|^{1-d}$ for some finite $c$ (because of the homogeneity of $K$), a simple computation shows that this definition is well-posed for a.e. $x$ and $K * \mu$ belongs to $L^p_{\mathrm{loc}}(\mathbb{R}^d)$ for every $p$ with $1 \leq p < \gamma(1)$.

**3.3. A second convolution operator.** Since $\nabla K$ is not summable on any neighbourhood of 0 (because it is homogeneous of degree $-d$), we cannot define $\nabla K * \mu$ by the usual integral formula. However, a classical result by A. P. Calderón and A. Zygmund shows that the convolution $\nabla K * \mu$ is well-defined at almost every point as a singular integral. More precisely, given the truncated kernels

$$(\nabla K)_\varepsilon(x) := \begin{cases} \nabla K(x) & \text{if } |x| \geq \varepsilon, \\ 0 & \text{if } |x| < \varepsilon, \end{cases} \tag{3.3}$$

then the functions $(\nabla K)_\varepsilon * \mu$ converge almost everywhere as $\varepsilon \to 0$ to a limit function which we denote by $\nabla K * \mu$. Moreover the following weak $L^1$-estimate holds:

$$\mathscr{L}^d\big(\{x : |\nabla K * \mu(x)| \geq t\}\big) \leq \frac{c\,\|\mu\|}{t} \quad \text{for every } t > 0, \tag{3.4}$$

where $c$ is a finite constant that depends only on $d$ and $K$.

If $\mu$ is replaced by a function in $L^1$, this statement follows, for instance, from Theorem 4 in Chapter II of [11];[10] extending that theorem to bounded measures requires only minor modifications in the proof.

We can now state the main result of this section.

**Theorem 3.4.** *Take $u := K * \mu$ as in §3.2. Then*

(i) *$u$ is $L^p$-differentiable for every $p$ with $1 \leq p < \gamma(1)$ and almost every $x \in \mathbb{R}^d$;*

(ii) *the derivative of $u$ is given by*

$$\nabla u = \nabla K * \mu + \beta_K f \quad \text{a.e.}, \tag{3.5}$$

*where $\nabla K * \mu$ is given in §3.3, $f$ is the Radon–Nikodym density of $\mu$ with respect to the Lebesgue measure*[11]*, and $\beta_K$ is the vector defined by*

$$\beta_K := \int_{S^{d-1}} x\, K(x)\, d\mathscr{H}^{d-1}(x)\,. \tag{3.6}$$

**Remark 3.5.** The range of $p$ for which $L^p$-differentiability holds is optimal. Take indeed $K(x) := |x|^{1-d}$ and

$$\mu := \sum_i 2^{-i}\delta_i\,,$$

where $\delta_i$ is the Dirac mass at $x_i$, and the set $\{x_i\}$ is dense in $\mathbb{R}^d$.

---

[10] In order to apply such theorem, the key point is that $\nabla K$ is of class $C^1$, homogeneous of degree $-d$, and satisfies the cancellation property (3.1).

[11] That is, the function $f$ such that the absolutely continuous part of $\mu$ with respect to $\mathscr{L}^d$ can be written as $f \cdot \mathscr{L}^d$.

Since $K(x)$ does not belong to $L^{\gamma(1)}(U)$ for any neighbourhood $U$ of 0, the function $u := K * \mu$ does not belong to $L^{\gamma(1)}(U)$ for any open set $U$ in $\mathbb{R}^d$. Hence $u$ does not belong to $T^{0,\gamma(1)}(x)$ and consequently not even to $t^{1,\gamma(1)}(x)$ for every $x \in \mathbb{R}^d$.

Note that the previous construction works *as is* for any nontrivial positive kernel $K$; a suitable refinement allows to remove the positivity constraint.

The rest of this section is devoted to the proof of Theorem 3.4.

The key point is to show that $u$ is in $T^{1,p}(x)$ for all $x$ in some "large" set (Lemma 3.11). To achieve this, the basic strategy is quite standard, and consists in writing $u$ as sum of two functions $u_g$ and $u_b$ given by a suitable Calderón–Zygmund decomposition of the measure $\mu$. Then we use Lemma 3.9 to show that $u_g$ is a function of class $W^{1,q}$ for every $q \geq 1$, and therefore its differentiability is a well-established fact, and use Lemma 3.10 to estimate the derivative of $u_b$ on a large set. The latter lemma is the heart of the proof.

In the next three paragraphs we recall some classical tools of the theory of singular integrals.

**3.6. Singular integrals: the $L^q$ case.** We have seen in §3.3 that the convolution product $\nabla K * \mu$ is well-defined at almost every point as a singular integral.

When $\mu$ is replaced by a function $f$ in $L^q(\mathbb{R}^d)$ with $1 < q < \infty$ there holds more: taking $(\nabla K)_\varepsilon$ as in (3.3), then $\|(\nabla K)_\varepsilon * f\|_q \leq c\|f\|_q$ for every $\varepsilon > 0$ and every $f \in L^q(\mathbb{R}^d)$, where $c$ is a finite constant that depends only on $K$ and $q$. Moreover, as $\varepsilon$ tends to 0, the functions $(\nabla K)_\varepsilon * f$ converges in the $L^q$-norm to some limit that we denote by $\nabla K * f$; in particular $f \mapsto \nabla K * f$ is a bounded linear operator from $L^q(\mathbb{R}^d)$ into $L^q(\mathbb{R}^d)$.

These statements follow from Theorem 3 in Chapter II of [11].

**3.7. Marcinkiewicz integral.** Let $\mu$ be a bounded (possibly vector-valued) measure on $\mathbb{R}^d$, and $F$ a closed set in $\mathbb{R}^d$. Then the Marcinkiewicz integral

$$I(\mu, F, x) := \int_{\mathbb{R}^d \setminus F} \frac{\operatorname{dist}(y, F)}{|x-y|^{d+1}} \, d|\mu|(y) \tag{3.7}$$

is finite for almost every $x \in F$, and more precisely

$$\int_F I(\mu, F, x) \, dx \leq c \, \|\mu\| \, , \tag{3.8}$$

where $c$ is a finite constant that depends only on $d$. This is a standard estimate, see §2.3 in Chapter I of [11].

**3.8. Maximal function.** Let $\mu$ be a bounded (possibly vector-valued) measure on $\mathbb{R}^d$. The maximal function associated to $\mu$ is

$$M(\mu, x) := \sup_{\rho > 0} \frac{|\mu|(B(x,\rho))}{\omega_d \, \rho^d} \, . \tag{3.9}$$

Then $M(\mu,x)$ is finite for almost every $x$, and more precisely the following weak $L^1$-estimate holds:

$$\mathscr{L}^d(\{x : M(\mu,x) \geq t\}) \leq \frac{c\,\|\mu\|}{t} \quad \text{for every } t>0, \tag{3.10}$$

where $c$ is a finite constant that depends only on $d$.

In case $\mu$ is absolutely continuous with respect to the Lebesgue measure this statement can be found in §1.3 in Chapter I of [11]; the proof for a general measure is essentially the same, cf. §4.1 in Chapter III of [11].

**Lemma 3.9.** *Let $f$ be a function in $L^1 \cap L^q(\mathbb{R}^d)$ for some $q$ with $1 < q < +\infty$, and let $u := K * f$. Then $u$ belongs to $L^1_{\rm loc}(\mathbb{R}^d)$ and the distributional derivative of $u$ is given by*

$$\nabla u = \nabla K * f + \beta_K f \tag{3.11}$$

*where $\nabla K * f$ is defined in §3.6, and $\beta_K$ is given in (3.6).*

*Since $\nabla K * f$ belongs to $L^q(\mathbb{R}^d)$, then $\nabla u$ belongs to $L^q(\mathbb{R}^d)$, and therefore $u$ is $L^{\gamma(q)}$-differentiable almost everywhere when $q < d$, and is continuous and differentiable almost everywhere in the classical sense when $q > d$ (in both cases the pointwise derivative agrees with the distributional one almost everywhere).*

*Proof.* Note that the second part of the statement is easily derived from §3.6 and formula (3.11) using the standard differentiability result for Sobolev functions (see for instance Sections 6.1.2 and 6.2 of [6]) and the fact that $K * f$ is continuous when $q > d$ (a matter of elementary estimates).

It remains to prove formula (3.11). For every $\varepsilon > 0$ consider the truncated kernel $K_\varepsilon$ defined as in (3.3), that is, $K_\varepsilon := 1_{\mathbb{R}^d \setminus B(\varepsilon)} K$. Then the distributional derivative of $K_\varepsilon$ is given by

$$\nabla K_\varepsilon = (\nabla K)_\varepsilon + \sigma_\varepsilon$$

where $\sigma_\varepsilon$ is the (vector-valued) measure given by the restriction of the Hausdorff measure $\mathscr{H}^{d-1}$ to the sphere $\partial B(\varepsilon)$ multiplied by the vector field $K(x)\,x/|x|$. Hence

$$\nabla(K_\varepsilon * f) = (\nabla K)_\varepsilon * f + \sigma_\varepsilon * f\,,$$

and we obtain (3.11) by passing to the limit as $\varepsilon \to 0$ in this equation.

In doing so we use the following facts:

- (i) $K_\varepsilon \to K$ in the $L^1$-norm, and therefore $\nabla(K_\varepsilon * f) \to \nabla(K * f) = \nabla u$ in the sense of distributions;
- (ii) $(\nabla K)_\varepsilon * f \to \nabla K * f$ in the $L^q$-norm (see §3.6);
- (iii) the measures $\sigma_\varepsilon$ converge in the sense of measures to $\beta_K$ times the Dirac mass at 0, and then $\sigma_\varepsilon * f \to \beta_K f$ in the $L^q$-norm. □

**Lemma 3.10.** *Let $F$ be a closed set in $\mathbb{R}^d$, $\{E_i\}$ a countable family of pairwise disjoint sets in $\mathbb{R}^d$ which do not intersect $F$, and $\mu$ a bounded real-valued measure on $\mathbb{R}^d$ such that*

- (i) *$|\mu|(\mathbb{R}^d \setminus \cup_i E_i) = 0$, and in particular $|\mu|(F) = 0$;*

(ii) $\mu(E_i) = 0$ *for every* $i$;

(iii) *there exist finite and strictly positive constants* $c_1$ *and* $c_2$ *such that, for every* $i$,

$$c_1 \operatorname{dist}(F, \operatorname{conv}(E_i)) \le \operatorname{diam}(E_i) \le c_2 \operatorname{dist}(F, \operatorname{conv}(E_i)) .$$

*Then, for every* $x \in F$ *and every* $p$ *with* $1 \le p < \gamma(1)$, *the function* $u := K * \mu$ *satisfies*

$$\left[ \fint_{B(\rho)} |u(x+h) - u(x)|^p \, dh \right]^{1/p} \le \big[ M(\mu, x) + I(\mu, F, x) \big] \, c\rho \,, \tag{3.12}$$

*where* $I(\mu, F, x)$ *and* $M(\mu, x)$ *are given in* (3.7) *and* (3.9), *respectively, and* $c$ *is a finite constant that depends only on* $c_1$, $c_2$, $p$, $d$ *and* $K$.[12]

*Thus* $u$ *belongs to* $T^{1,p}(x)$ *for every* $x \in F$ *such that* $M(\mu, x)$ *and* $I(\mu, F, x)$ *are finite, that is, for almost every* $x \in F$.

*Proof.* We fix a point $x \in F$ and $\rho > 0$, and denote by $J$ the set of all indexes $i$ such that $\operatorname{dist}(x, \operatorname{conv}(E_i)) < 2\rho$.

Using assumption (i) we decompose $u$ as

$$u = \sum_i u_i \,, \tag{3.13}$$

where $u_i := K * \mu_i$ and $\mu_i$ is the restriction of the measure $\mu$ to the set $E_i$.

*Step* 1. *Estimate of* $|u_i(x)|$ *for* $i \in J$. Choose an arbitrary point $y_i \in E_i$, and for every $s \in [0, 1]$ set

$$g(s) := \int_{E_i} K(x - (sy + (1-s)y_i)) \, d\mu(y) \,.$$

Then[13]

$$u_i(x) = \int_{E_i} K(x-y) \, d\mu(y) = \int_{E_i} K(x-y) - K(x-y_i) \, d\mu(y) = g(1) - g(0) \,,$$

and by applying the mean value theorem to the function $g$ we obtain that there exists $s \in [0, 1]$ such that $u_i(x) = g(1) - g(0) = \dot g(s)$, that is,

$$u_i(x) = \int_{E_i} \nabla K \big( \underbrace{x - (sy + (1-s)y_i)}_{z} \big) \, (y_i - y) \, d\mu(y) \,. \tag{3.14}$$

[12] When we apply this lemma later on, the constants $c_1$ and $c_2$ will depend only on $d$, and therefore the constant $c$ in (3.12) will depend only on $p$, $d$ and $K$.

[13] The second identity follows from the fact that $\mu(E_i) = 0$ by assumption (ii), and the third one follows from the definition of $g$.

Since $\nabla K$ is homogeneous of degree $-d$, there holds[14] $|\nabla K(z)| \leq c|z|^{-d}$,and taking into account that $|z| \geq \mathrm{dist}(x, \mathrm{conv}(E_i))$ and $\mathrm{dist}(x, \mathrm{conv}(E_i)) < 2\rho$ we get

$$\begin{aligned} |\nabla K(z) \cdot (y_i - y)| &\leq |\nabla K(z)|\, |y_i - y| \\ &\leq \frac{c\,\mathrm{diam}(E_i)}{\mathrm{dist}(x, \mathrm{conv}(E_i))^d} \leq \frac{c\rho\,\mathrm{diam}(E_i)}{\mathrm{dist}(x, \mathrm{conv}(E_i))^{d+1}} \,. \end{aligned}$$

Moreover, for every $y \in E_i$ assumption (iii) implies $\mathrm{diam}(E_i) \leq c\,\mathrm{dist}(y, F)$ and $|x - y| \leq c\,\mathrm{dist}(x, \mathrm{conv}(E_i))$, and therefore

$$|\nabla K(z) \cdot (y_i - y)| \leq \frac{c\rho\,\mathrm{dist}(y, F)}{|x - y|^{d+1}} \,.$$

Plugging the last estimate in (3.14) we obtain

$$|u_i(x)| \leq c\rho \int_{E_i} \frac{\mathrm{dist}(y, F)}{|x - y|^{d+1}}\, d|\mu|(y) \,. \tag{3.15}$$

*Step* 2. *Estimate of* $|u_i(x + h)|$ *for* $i \in J$. We take $p$ with $1 \leq p < \gamma(1)$ and denote by $p'$ the conjugate exponent of $p$, that is, $1/p' + 1/p = 1$. We also choose a positive test function $\varphi$ on $B(\rho)$, and denote by $\|\varphi\|_{p'}$ the $L^{p'}$-norm of $\varphi$ with respect to the Lebesgue measure on $B(\rho)$ normalized to a probability measure. Then[15]

$$\begin{aligned} &\fint_{B(\rho)} |u_i(x + h)|\, \varphi(h)\, dh \\ &\quad\leq \int_{E_i} \Big[ \fint_{B(\rho)} |K(x + h - y)|\, \varphi(h)\, dh \Big]\, d|\mu|(y) \\ &\quad\leq \int_{E_i} \Big[ \fint_{B(\rho)} |\varphi(h)|^{p'}\, dh \Big]^{1/p'} \Big[ \fint_{B(\rho)} |K(x + h - y)|^p dh \Big]^{1/p} d|\mu|(y) \\ &\quad\leq \int_{E_i} \|\varphi\|_{p'} \Big[ \frac{c}{\rho^d} \int_{B(\rho)} \frac{dh}{|x + h - y|^{p(d-1)}} \Big]^{1/p} d|\mu|(y) \\ &\quad\leq \frac{c}{\rho^{d/p}} \|\varphi\|_{p'} \Big[ \int_{B(c\rho)} \frac{dz}{|z|^{p(d-1)}} \Big]^{1/p} |\mu|(E_i) \leq \frac{c}{\rho^{d-1}} \|\varphi\|_{p'}\, |\mu|(E_i) \,, \end{aligned}$$

[14] Here and in the rest of this proof we use the letter $c$ to denote any finite and strictly positive constant that depends only on $c_1$, $c_2$, $p$, $d$, and $K$. Accordingly, the value of $c$ may change at every occurrence.

[15] For the first inequality we use the definition of $u_i$ and Fubini's theorem; for the second one we use Hölder inequality, for the third one we use that $K$ is homogeneous of degree $1 - d$ and therefore $|K(x)| \leq c|x|^{1-d}$; for the fourth one we use the change of variable $z = x + h - y$ and the fact that for every $y \in E_i$ assumption (iii) yields

$$|x + h - y| \leq |x - y| + |h| \leq \mathrm{dist}(x, \mathrm{conv}(E_i)) + \mathrm{diam}(\mathrm{conv}(E_i)) + \rho \leq c\rho \,.$$

Note that the integral in the last line is finite if and only if $p < \gamma(1)$. Here is the only place in the entire proof where this upper bound on $p$ is needed.

and taking the supremum over all test function $\varphi$ with $\|\varphi\|_{p'} \le 1$ we finally get

$$\Big[\fint_{B(\rho)} |u_i(x+h)|^p\, dh\Big]^{1/p} \le \frac{c}{\rho^{d-1}}\, |\mu|(E_i)\,. \tag{3.16}$$

*Step* 3. Using the estimates (3.15) and (3.16), and taking into account that $E_i$ is contained in $B(x, c\rho)$ for every $i \in J$ (use assumption (iii) as in Footnote 15) and that $|\mu|(F) = 0$ (recall assumption (i)), we get

$$\begin{aligned}
&\sum_{i\in J}\Big[\fint_{B(\rho)} |u_i(x+h) - u_i(x)|^p\, dh\Big]^{1/p} \\
&\qquad \le \sum_{i\in J}\Big[\fint_{B(\rho)} |u_i(x+h)|^p\, dh\Big]^{1/p} + \sum_{i\in J}|u_i(x)| \\
&\qquad \le c\frac{|\mu|(B(x,c\rho))}{\rho^{d-1}} + c\rho\int_{B(x,c\rho)} \frac{\operatorname{dist}(y,F)}{|x-y|^{d+1}}\, d|\mu|(y) \le \big[M(\mu,x) + I(\mu,F,x)\big]\, c\rho. 
\end{aligned} \tag{3.17}$$

*Step* 4. *Estimate of* $|u_i(x+h) - u_i(x)|$ *for* $i \notin J$. Let $y_i$ be a point in $E_i$. Then for every $h \in B(\rho)$ there exist $t, s \in [0,1]$ such that [16]

$$\begin{aligned}
u_i(x+h) - u_i(x) &= \int_{E_i} K(x+h-y) - K(x-y)\, d\mu(y) \\
&= \int_{E_i} \nabla K(x+th-y)\cdot h\, d\mu(y) \\
&= \int_{E_i} \big[\nabla K(x+th-y) - \nabla K(x+th-y_i)\big]\cdot h\, d\mu(y) \\
&= \int_{E_i} \big[\nabla^2 K\big(\underbrace{x+th-(sy+(1-s)y_i)}_{z}\big)\,(y_i - y)\big]\cdot h\, d\mu(y)\,.
\end{aligned} \tag{3.18}$$

Now, the fact that $\operatorname{dist}(x, \operatorname{conv}(E_i)) \ge 2\rho$ yield

$$\begin{aligned}
|z| &\ge |x - (sy + (1-s)y_i)| - t|h| \\
&\ge \operatorname{dist}(x, \operatorname{conv}(E_i)) - \rho \ge \frac{1}{2}\operatorname{dist}(x, \operatorname{conv}(E_i))\,,
\end{aligned}$$

and then, taking into account that $\nabla^2 K$ is homogeneous of degree $-d-1$,

$$\big|[\nabla^2 K(z)(y_i - y)]\cdot h\big| \le |\nabla^2 K(z)|\,|y_i - y|\,|h| \le \frac{c\operatorname{diam}(E_i)\,\rho}{\operatorname{dist}(x,\operatorname{conv}(E_i))^{d+1}}\,.$$

Moreover assumption (iii) implies that $\operatorname{diam}(E_i) \le c\operatorname{dist}(y, F)$ and $|x - y| \le c\operatorname{dist}(x, \operatorname{conv}(E_i))$ for every $y \in E_i$, and then

$$\big|[\nabla^2 K(z)(y_i - y)]\cdot h\big| \le \frac{c\operatorname{dist}(y,F)\,\rho}{|x-y|^{d+1}}\,.$$

[16] The second and fourth identities are obtained by applying the mean-value theorem as in Step 1, the third one follows from assumption (ii).

Hence (3.18) yields

$$|u_i(x+h) - u_i(x)| \le c\rho \int_{E_i} \frac{\operatorname{dist}(y,F)}{|x-y|^{d+1}}\, d|\mu|(y)\,. \tag{3.19}$$

*Step* 5. Inequality (3.19) yields

$$\sum_{i\notin J} \Big[ \fint_{B(\rho)} |u_i(x+h) - u_i(x)|^p \, dh \Big]^{1/p} \le I(\mu, F, x)\, c\rho\,,$$

and recalling estimate (3.17) and formula (3.13) we finally obtain (3.12). □

**Lemma 3.11.** *Take $u$ as in Theorem 3.4. Take $t>0$ and let*

$$F_t := \{x \in \mathbb{R}^d : \ M(\mu, x) \le t\}\,,$$

*where $M(\mu,x)$ is the maximal function defined in (3.9).*

*Then $u$ belongs to $T^{1,p}(x)$ for every $p$ with $1 \le p < \gamma(1)$ and almost every $x \in F_t$.*

*Proof. Step* 1. *Calderón–Zygmund decomposition of $\mu$ and $u$.* Since $M(\mu,x)$ is lower semicontinuous at $x$ (being the supremum of a family of lower semicontinuous functions), the set $F_t$ is closed.

We take a Whitney decomposition of the open set $\mathbb{R}^d \setminus F_t$, that is, a sequence of closed cubes $Q_i$ with pairwise disjoint interiors such that the union of all $Q_i$ is $\mathbb{R}^d \setminus F_t$, and the distance between $F_t$ and each $Q_i$ is comparable to the diameter of $Q_i$, namely

$$c_1 \operatorname{dist}(F_t, Q_i) \le \operatorname{diam}(Q_i) \le c_2 \operatorname{dist}(F_t, Q_i)\,, \tag{3.20}$$

where $c_1$ and $c_2$ depend only on $d$ (see §3.1 in Chapter I of [11]).

We consider now the sets $E_i$ obtained by removing from each $Q_i$ part of its boundary so that the sets $E_i$ are pairwise disjoint and still cover $\mathbb{R}^d \setminus F_t$.

The Calderón–Zygmund decomposition of $\mu$ is $\mu = \mu_g + \mu_b$, where the "good" part $\mu_g$ is defined by

$$\mu_g := 1_{F_t} \cdot \mu + \sum_i a_i 1_{E_i} \cdot \mathscr{L}^d \quad \text{with } a_i := \frac{\mu(E_i)}{\mathscr{L}^d(E_i)}\,, \tag{3.21}$$

and the "bad" part $\mu_b$ is

$$\mu_b := \sum_i 1_{E_i} \cdot \mu - a_i 1_{E_i} \cdot \mathscr{L}^d\,. \tag{3.22}$$

From this definition and that of $a_i$ we obtain

$$\|\mu_b\| \le \sum_i 2|\mu|(E_i) = 2|\mu|(\mathbb{R}^d \setminus F_t)\,. \tag{3.23}$$

Next we decompose $u$ as

$$u = u_g + u_b\,,$$

where $u_g := K * \mu_g$ and $u_b := K * \mu_b$. To conclude the proof we need to show that $u_g$ and $u_b$ belong to $T^{1,p}(x)$ for every $1 \le p < \gamma(1)$ and almost every $x \in F_t$. This will be done in the next steps.

*Step* 2. *The measure $\mu_g$ can be written as $g\cdot\mathscr{L}^d$, with $g\in L^\infty(\mathbb{R}^d)$.* It suffices to show that

(i) the measure $1_{F_t}\cdot\mu$ can be written as $\tilde g\cdot\mathscr{L}^d$ with $\tilde g\in L^\infty(\mathbb{R}^d)$;

(ii) the number $a_i$ in (3.21) satisfy $|a_i|\le ct$ for some finite constant $c$ depending only on $d$.

Claim (i) follows by the fact that the Radon–Nikodym density of $|\mu|$ with respect to $\mathscr{L}^d$ is bounded by $t$ at *every* point $x$ of $F_t$, because $M(\mu,x)\le t$.

To prove claim (ii), note that each $E_i$ is contained in $Q_i$, which in turn is contained in a ball $B(x_i,r_i)$ for a suitable $x_i\in F_t$ and $r_i$ comparable to the diameter of $Q_i$ (use (3.20)). Hence, taking into account that $M(\mu,x_i)\le t$,

$$|\mu|(E_i)\le|\mu|(B(x_i,r_i))\le t\,\mathscr{L}^d(B(x_i,r_i))\le ct\,\mathscr{L}^d(Q_i)=ct\,\mathscr{L}^d(E_i)\,,$$

and this implies $|a_i|\le ct$.

*Step* 3. *$u_g$ is differentiable at $x$ (and in particular belongs to $T^{1,p}(x)$ for every $1\le p<+\infty$) for almost every $x\in\mathbb{R}^d$.* Since the measure $\mu_g$ is bounded, the function $g$ in Step 2 belongs to $L^1\cap L^\infty(\mathbb{R}^d)$. Then, by interpolation, $g$ belongs to $L^1\cap L^q(\mathbb{R}^d)$ for any $q>d$, and Lemma 3.9 implies that $u_g=K*g$ is differentiable almost everywhere.

*Step* 4. *$u_b$ belongs to $T^{1,p}(x)$ for almost every $x\in F_t$ and every $1\le p<\gamma(1)$.* It suffices to apply Lemma 3.10 to the set $F_t$, the measure $\mu_b$, and the sets $E_i$ (use equations (3.20), (3.21) and (3.22) to check that the assumptions of that lemma are verified). □

*Proof of Theorem* 3.4. (i) It suffices to apply Lemma 3.11 and Remark 2.3 (iii), and take into account that the sets $F_t$ form an increasing family whose union cover almost all of $\mathbb{R}^d$ (because the maximal function $M(\mu,x)$ is finite almost everywhere).

(ii) Since we already know that $u$ is $L^p$-differentiable almost everywhere, we have only to prove identity (3.5).

Moreover, by the argument used in the proof of statement (i) above, it suffices to show that (3.5) holds almost everywhere in the set $F_t$ defined in Lemma 3.11 for every given $t>0$.

*Step* 1. *Decomposition of $\mu$ and $u$.* We fix for the time being $\varepsilon>0$, and choose a closed set $C$ contained in $\mathbb{R}^d\setminus F_t$ such that $|\mu|(\mathbb{R}^d\setminus(F_t\cup C))\le\varepsilon$.

We decompose $\mu$ as $\mu'+\mu''$ where $\mu'$ and $\mu''$ are the restrictions of $\mu$ to the sets $\mathbb{R}^d\setminus C$ and $C$, respectively, and then we further decompose $\mu'$ as $\mu'_g+\mu'_b$ as in the proof of Lemma 3.11. Thus $\mu=\mu'_g+\mu'_b+\mu''$, and accordingly we decompose $u$ as

$$u=u'_g+u'_b+u''$$

where $u'_g:=K*\mu'_g$, $u'_b:=K*\mu'_b$, and $u'':=K*\mu''$.

Using estimate (3.23) and taking into account the definition of $\mu'$ and the choice of $C$ we get

$$\|\mu'_b\|\le 2\,|\mu'|(\mathbb{R}^d\setminus F_t)=2\,|\mu|(\mathbb{R}^d\setminus(F_t\cup C))\le 2\varepsilon\,. \tag{3.24}$$

*Step* 2. *Derivative of $u'_g$.* Going back to the proof of Lemma 3.11, we see that $\mu'_g$ can be written as $g \cdot \mathscr{L}^d$ with $g \in L^1 \cap L^\infty(\mathbb{R}^d)$, and therefore $u'_g := K * g$ is differentiable almost everywhere. Moreover formula (3.11) yields

$$\nabla u'_g = \nabla K * g + \beta_K g = \nabla K * \mu'_g + \beta_K g \quad \text{a.e.}$$

Next we note that the restrictions of the measures $\mu'_g$, $\mu'$ and $\mu$ to the set $F_t$ agree, and therefore $g = f$ almost everywhere on $F_t$, where $f$ is the Radon–Nikodym density of $\mu$ with respect to $\mathscr{L}^d$. Thus the previous identity becomes

$$\nabla u'_g = \nabla K * \mu'_g + \beta_K f \quad \text{a.e. in } F_t. \tag{3.25}$$

*Step* 3. *Derivative of $u'_b$.* Going back to the proof of Lemma 3.11 we see that we can use Lemma 3.10 to show that $u'_b$ belongs to $T^{1,1}(x)$ for almost every $x \in F_t$. By Remark 2.3 (iii) we have that $u'_b$ is $L^1$-differentiable almost everywhere in $F_t$, and therefore estimate (3.12) in Lemma 3.10 yields

$$|\nabla u'_b(x)| \leq c\, M(\mu'_b, x) + c\, I(\mu'_b, F_t, x) \quad \text{for a.e. } x \in F_t.\text{[17]} \tag{3.26}$$

*Step* 4. *Derivative of $u''$.* By construction, the support of the measure $\mu''$ is contained in the closed set $C$ and therefore the convolution $u'' := K * \mu''$ can be defined in the classical sense, and is smooth, at every point of the open set $\mathbb{R}^d \setminus C$, which contains $F_t$. Hence

$$\nabla u'' = \nabla K * \mu'' \quad \text{everywhere in } F_t.\text{[18]} \tag{3.27}$$

*Step* 5. Putting together equations (3.25) and (3.27), and the fact that $\mu = \mu'_g + \mu'_b + \mu''$, we obtain

$$\nabla u - (\nabla K * \mu + \beta_K f) = \nabla u'_b - \nabla K * \mu'_b \quad \text{a.e. in } F_t,$$

and using estimate (3.26),

$$\begin{aligned} &|\nabla u - (\nabla K * \mu + \beta_K f)| \\ &\qquad \leq c\, M(\mu'_b, \cdot) + c\, I(\mu'_b, F_t, \cdot) + |\nabla K * \mu'_b| \quad \text{a.e. in } F_t. \end{aligned} \tag{3.28}$$

Finally, using the fact that $\|\mu'_b\| \leq 2\varepsilon$ (see (3.24)) and estimates (3.10), (3.8), and (3.4), we obtain that each term at the right-hand side of (3.28) is smaller than $\sqrt{\varepsilon}$ outside an exceptional set with measure at most $c\sqrt{\varepsilon}$.

Since $\varepsilon$ is arbitrary, we deduce that $\nabla u = \nabla K * \mu + \beta_K f$ almost everywhere in $F_t$, and the proof is complete. □

[17] Here and in the rest of this proof we use the letter $c$ to denote any finite and strictly positive constant that depends only on $d$ and $K$.

[18] We tacitly use that this "classical" convolution agrees (a.e.) with the singular integral.

## 4. Further differentiability results

**4.1. Additional kernels.** Let $G : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ be the fundamental solution of the Laplacian $(-\Delta)$ on $\mathbb{R}^d$, that is,

$$G(x) := \begin{cases} \dfrac{1}{d(d-2)\omega_d}\,|x|^{2-d} & \text{if } d > 2, \\ -\dfrac{1}{2\pi}\log|x| & \text{if } d = 2, \end{cases}$$

and, for every $h = 1, \dots, d$, let

$$K_h(x) := -\partial_h G(x) = \frac{1}{d\omega_d}|x|^{-d}x_h\,.$$

We can now state the main results of this section; proofs will be given after Remark 4.5.

**Proposition 4.2.** *Let $v = (v_1, \dots, v_d)$ be a vector field in $L^1(\mathbb{R}^d)$ whose distributional curl and divergence are bounded measures, and denote by $\mu_0$ and $\mu_{hk}$ the measures*

$$\mu_0 := \operatorname{div} v \quad \textit{and} \quad \mu_{hk} := (\operatorname{curl} v)_{hk} = \partial_h v_k - \partial_k v_h \tag{4.1}$$

*for every $1 \le h, k \le d$. Then, for every $k = 1, \dots, d$, there holds*

$$v_k = K_k * \mu_0 + \sum_{h=1}^{d} K_h * \mu_{hk} \quad \textit{a.e.} \tag{4.2}$$

*Therefore each $v_k$, and consequently also $v$, is $L^p$-differentiable at almost every $x \in \mathbb{R}^d$ for every $p$ with $1 \le p < \gamma(1)$.*

**Proposition 4.3.** *Let $v$ be a vector field in $L^1(\mathbb{R}^d)$ with bounded deformation, that is, the distributional symmetric derivative $\frac{1}{2}(\nabla v + \nabla^t v)$ is a bounded measure, and denote by $\lambda_{hk}$ the measures*

$$\lambda_{hk} := \frac{1}{2}(\partial_h v_k + \partial_k v_h) \tag{4.3}$$

*for every $1 \le h, k \le d$. Then for every $k = 1, \dots, d$ there holds*

$$v_k = \sum_{h=1}^{d} \big(2K_h * \lambda_{hk} - K_k * \lambda_{hh}\big) \quad \textit{a.e.} \tag{4.4}$$

*Therefore each $v_k$, and consequently also $v$, is $L^p$-differentiable at almost every $x \in \mathbb{R}^d$ for every $p$ with $1 \le p < \gamma(1)$.*

**Proposition 4.4.** *Let $\Omega$ be an open set in $\mathbb{R}^d$, and $w$ a real function in $L^1_{\mathrm{loc}}(\Omega)$ whose distributional Laplacian is a locally bounded measure. Then $w$ admits an $L^p$-Taylor expansion of order two for a.e. $x \in \mathbb{R}^d$ and every $1 \le p < \gamma(\gamma(1))$.*

*In particular $w$ has the Lusin property with functions of class $C^2$.*

**Remark 4.5.** (i) Using statement (ii) in Theorem 3.4 one can write an explicit formula for the (pointwise) derivatives of the vector fields $v$ considered in Propositions 4.2 and 4.3.

(ii) Let $\Omega$ be any open set in $\mathbb{R}^d$. The differentiability property stated in Proposition 4.2 holds also for vector fields $v$ in $L^1_{\mathrm{loc}}(\Omega)$ whose curl and divergence are locally bounded measures. The key observation is that given a smooth cutoff function $\varphi$ on $\mathbb{R}^d$ with support contained in $\Omega$, then $\varphi v$ is a vector field in $L^1(\mathbb{R}^d)$ and its curl and divergence are bounded measures.

The same argument applies to Proposition 4.3.

(iii) The range of $p$ in Proposition 4.4 is optimal, and this can be shown by taking $\Omega := \mathbb{R}^d$ and $w := G * \mu$, where $G$ is given in §4.1 and $\mu$ is given in Remark 3.5. Indeed $-\Delta w = \mu$ and one easily checks that $w$ does not belong to $L^{\gamma(\gamma(1))}(U)$ for any open set $U$ in $\mathbb{R}^d$. Hence $w$ does not belong to $T^{0,\gamma(\gamma(1))}(x)$ for any $x \in \mathbb{R}^d$, and consequently not even to $t^{2,\gamma(\gamma(1))}(x)$.

(iv) The range of $p$ in Proposition 4.2 is also optimal. Let indeed $v := \nabla w$ where $w$ is the function constructed above: then the curl of $w$ vanishes and the divergence agrees with the measure $-\mu$, and $v$ does not belong to $L^{\gamma(1)}(U)$ for any open set $U$ in $\mathbb{R}^d$ (otherwise the Sobolev embedding would imply that $w$ belongs to $L^{\gamma(\gamma(1))}(U)$).

(v) We do not know if the range of $p$ in Proposition 4.3 is optimal, and more precisely whether or not a map in $BD$ belongs to $t^{1,\gamma(1)}(x)$ for almost every $x$. Note that the argument used in points (iii) and (iv) above does not apply here because the space $BD$ does embed in $L^{\gamma(1)}$ for regular domains, see Proposition 1.2 in [12].

*Proof of Proposition* 4.2. By applying the Fourier transform to the identities in (4.1) we obtain

$$\sum_h i\xi_h\,\hat v_h = \hat\mu_0 \quad\text{and}\quad i\xi_h\,\hat v_k = i\xi_k\,\hat v_h + \hat\mu_{hk}\,, \tag{4.5}$$

where $i = \sqrt{-1}$ and $\xi$ denotes the Fourier variable.

We multiply the second identity in (4.5) by $-i\xi_h$ and sum over all $h$; taking into account the first identity in (4.5) we get

$$|\xi|^2\hat v_k = \xi_k \sum_h \xi_h \hat v_h - \sum_h i\xi_h\,\hat\mu_{hk} = -i\xi_k\,\hat\mu_0 - \sum_h i\xi_h\,\hat\mu_{hk}\,.$$

Now $-\Delta G = \delta_0$ implies $\hat G = |\xi|^{-2}$ and then $\hat K_h = -i\xi_h \hat G = -i\xi_h|\xi|^{-2}$ (see §4.1). Thus the previous identity yields

$$\hat v_k = \frac{-i\xi_k}{|\xi|^2}\,\hat\mu_0 + \sum_h \frac{-i\xi_h}{|\xi|^2}\,\hat\mu_{hk} = \hat K_k\,\hat\mu_0 + \sum_h \hat K_h\,\hat\mu_{hk}\,,$$

and (4.2) follows by taking the inverse Fourier transform. The rest of Proposition 4.2 follows from Theorem 3.4. □

Proposition 4.3 can be proved in the same way as Proposition 4.2; we omit the details.

**Lemma 4.6.** *Let $k \geq 0$ be an integer, and $p \geq 1$ a real number. Let $u$ be a function in $W^{1,1}(\Omega)$ where $\Omega$ is a bounded open set in $\mathbb{R}^d$, and assume that the distributional derivative $\nabla u$ belongs to $t^{k,p}(x)$ (respectively, $T^{k,p}(x)$) for some point $x \in \Omega$. Then $u$ belongs to $t^{k+1,\gamma(p)}(x)$ (respectively, $T^{k+1,\gamma(p)}(x)$).*

This lemma is contained in Theorem 11 of [4], at least in the case $\Omega = \mathbb{R}^d$ and $u$ with compact support (recall Remark 2.3 (i)). Note that we can always reduce to this case by multiplying $u$ by suitable cutoff functions.

*Proof of Proposition 4.4.* Apply Proposition 4.2 to the vector field $\nabla w$ and then use Lemma 4.6 (and recall § 2.4). □

We conclude this section with a comment on the last proof.

**Remark 4.7.** The key ingredient in the proof of the Lusin property for the functions $w$ considered in Proposition 4.4 is the $L^p$-differentiability of $\nabla w$. Here we want to argue that the approximate differentiability of $\nabla w$ in the sense of Remark 2.3 (v) would have not been sufficient. In other words, Proposition 4.4 cannot be derived from the differentiability result in [9].

We claim indeed that even in dimension $d = 1$, the approximate differentiability of the derivative of a function $w$ at almost every point of a set $E$ is not enough to prove that $w$ has the Lusin property with functions of class $C^2$ on $E$. More precisely, there exists a function $w : \mathbb{R} \to \mathbb{R}$ of class $C^1$ such that $\dot{w} = 0$ on some set $E$ with positive measure (and therefore $\dot{w}$ is approximately differentiable with derivative equal to 0 at almost every point of $E$) but $w$ does not have the Lusin property with functions of class $C^2$ on $E$.

The construction of such a function is briefly sketched in the next paragraph.

**4.8. Example.** Fix $\lambda$ such that $1/4 < \lambda < 1/2$. For every $n = 0, 1, 2 \dots$ let $E_n$ be the union of the closed intervals $I_{n,k}$, $k = 1, \dots, 2^n$, obtained as follows: $I_{0,1} = E_0$ is a closed interval with length 2, and the intervals $I_{n+1,k}$ are obtained by removing from each $I_{n,k}$ a concentric open interval $J_{n,k}$ with length $(1-2\lambda)\lambda^n$.

Now let $E$ be the intersection of the sets $E_n$. This construction of Cantor type produces a compact set $E$ with empty interior such that $\mathscr{L}^1(E) = 1$.

Next we construct a non-negative continuous function $v : \mathbb{R} \to \mathbb{R}$ such that $v = 0$ outside the union of the intervals $J_{n,k}$ over all $n$ and $k$, and the integral of $v$ over each $J_{n,k}$ is equal to $(1-2\lambda)\lambda^n$.

Finally we take $w$ so that $\dot{w} = v$.

It is easy to verify that for every $n$ the set $E'_n := w(E_n)$ is the union of the disjoint intervals $I'_{n,k} := w(I_{n,k})$, $k = 1, \dots, 2^n$, and that these intervals have length $\lambda^n$. Moreover $E'_{n+1}$ can be written as the union of two disjoint copies of $E'_n$ scaled by a factor $\lambda$. Therefore the set $E' := w(E)$ can be written as the union of two disjoint copies of itself scaled by a factor $\lambda$. In other words, $E'$ is a self-similar fractal determined by two homoteties with scaling factor $\lambda$; it is then well known that $E'$ has Hausdorff dimension $s := \log 2/\log(1/\lambda)$ (see [7], Section 8.3).

Moreover, denoting by $\mu$ the push-forward according to $w$ of the restriction of the Lebesgue measure to $E$, one easily checks that $\mu$ is supported on $E'$ and satisfies $\mu(I'_{n,k}) = 2^{-n}$ for every $n$ and $k$. Therefore $\mu$ agrees with the canonical probability measure associated to the fractal $E'$, which in turn agrees, up to a constant factor, with the restriction of $\mathscr{H}^s$ to $E'$ (see [7], Section 8.3). In particular, since $s > 1/2$ (recall that $\lambda > 1/4$), we have that $\mu(A) = 0$ for every set $A$ which is $\sigma$-finite with respect to $\mathscr{H}^{1/2}$.

To show that $w$ does not have the Lusin property with functions of class $C^2$ on $E$ it is now sufficient to recall the following lemma.

**Lemma 4.9.** *Let $E$ be a set in $\mathbb{R}$ and $u : \mathbb{R} \to \mathbb{R}$ be a function of class $C^1$ such that $u$ has the Lusin property with functions of class $C^2$ on $E$ and $\dot{u} = 0$ on $E$. Then the push-forward of $1_E \cdot \mathscr{L}^1$ according to $w$ is supported on a set which is $\mathscr{H}^{1/2}$-negligible.*

*Proof.* It suffices to apply the definition of Lusin property and the fact that a function $u\colon \mathbb{R} \to \mathbb{R}$ of class $C^2$ maps any bounded set where $\dot{u} = 0$ into an $\mathscr{H}^{1/2}$-negligible set (this is a particular case of Sard's theorem, see for instance Theorem 3.4.3 in [8]). □

## References

[1] Alberti, G., Bianchini, S. and Crippa G.: A uniqueness result for the continuity equation in two dimensions. *J. Eur. Math. Soc.* **16** (2014), no. 2, 201–234.

[2] Ambrosio, L., Coscia, A. and Dal Maso, G.: Fine properties of functions with bounded deformation. *Arch. Rational Mech. Anal.* **139** (1997), no. 3, 201–238.

[3] Calderón, A. P. and Zygmund, A.: On the existence of certain singular integrals. *Acta Math.* **88** (1952), 85–139.

[4] Calderón, A. P. and Zygmund, A.: Local properties of solutions of elliptic partial differential equations. *Studia Math.* **20** (1961), 171–225.

[5] Conti, S., Faraco, D. and Maggi, F.: A new approach to counterexamples to $L^1$ estimates: Korn's inequality, geometric rigidity, and regularity for gradients of separately convex functions. *Arch. Ration. Mech. Anal.* **175** (2005), no. 2, 287–300.

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

[7] Falconer, K. J.: *The geometry of fractal sets.* Cambridge Tracts in Mathematics 85, Cambridge University Press, Cambridge, 1986.

[8] Federer, H.: *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften 153, Springer, Berlin-New York, 1969. Reprinted in the series Classics in Mathematics, Springer, Berlin-Heidelberg, 1996.

[9] Hajłasz, P.: On approximate differentiability of functions with bounded deformation. *Manuscripta Math.* **91** (1996), no. 1, 61–72.

[10] Ornstein, D. A.: A non-equality for differential operators in the $L_1$ norm. *Arch. Rational Mech. Anal.* **11** (1962), 40–49.

[11] Stein, E. M.: *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series 30, Princeton University Press, Princeton, 1970.

[12] Temam, R. and Strang, G.: Functions of bounded deformation. *Arch. Rational Mech. Anal.* **75** (1980/81), no. 1, 7–21.

[13] Ziemer, W. P.: *Weakly differentiable functions. Sobolev spaces and functions of bounded variation.* Graduate Texts in Mathematics 120, Springer, New York, 1989.

Received January 27, 2012; revised May 21, 2012.

Giovanni Alberti: Dipartimento di Matematica, Università di Pisa, largo Pontecorvo 5, 56127 Pisa, Italy.
E-mail: galberti1@dm.unipi.it

Stefano Bianchini: SISSA, via Bonomea 265, 34136 Trieste, Italy.
E-mail: bianchin@sissa.it

Gianluca Crippa: Departement Mathematik und Informatik, Universität Basel, Rheinsprung 21, CH-4051 Basel, Switzerland.
E-mail: gianluca.crippa@unibas.ch

This work has been partially supported by the Italian Ministry of Education, University and Research (MIUR) through the 2008 PRIN Grant "Trasporto ottimo di massa, disuguaglianze geometriche e funzionali e applicazioni" and by the European Research Council through the 2009 Starting Grant "Hyperbolic Systems of Conservation Laws: singular limits, properties of solutions and control problems".