Failure of the chain rule in the non steady two-dimensional setting

Stefano Bianchini and Paolo Bonicatto

Abstract In [CGSW17], the authors provide, via an abstract convex integration method, a vast class of counterexamples to the chain rule problem for the divergence operator applied to bounded, autonomous vector fields in $\mathbf{b}\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \geq 3$. By the analysis of [BG16] the assumption $d \geq 3$ is essential, as in the two dimensional setting, under the further assumption $\mathbf{b} \neq 0$ a.e., the Hamiltonian structure prevents from constructing renormalization defects.
In this note, following the ideas of [BBG16], we complete the analysis, by considering the non-steady, two dimensional case: we show that it is possible to construct a bounded, autonomous, divergence-free vector field $\mathbf{b}\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that there exists a non trivial, bounded distributional solution u to

$$\partial_t u + \operatorname{div}(u\mathbf{b}) = 0$$

for which the distribution $\partial_t \left(u^2\right) + \operatorname{div}\left(u^2\mathbf{b}\right)$ is *not* (representable by) a Radon measure.

MSC (2010): 35F05, 35A02, 35Q35

KEYWORDS: Chain Rule, Lipschitz maps, Transport and Continuity Equations, Disintegration

---

S. Bianchini
SISSA, via Bonomea 265, 34136 Trieste, Italy
e-mail: bianchin@sissa.it
P. Bonicatto
SISSA, via Bonomea 265, 34136 Trieste, Italy
e-mail: paolo.bonicatto@sissa.it

1 Introduction

In this paper we consider a variant of the classical problem of the chain rule for the divergence of a bounded vector field. Specifically, the classical problem of the chain rule reads as follows:

Problem 1. (Chain rule) Let $d \geq 2$ and assume that it is given a bounded, Borel vector field $\mathbf{b} \colon \mathbb{R}^d \to \mathbb{R}^d$, a bounded, scalar function $u \colon \mathbb{R}^d \to \mathbb{R}$ and Radon measures $\lambda, \mu \in \mathscr{M}(\mathbb{R}^d)$ such that

$$\operatorname{div} \mathbf{b} = \lambda, \tag{1a}$$

$$\operatorname{div}(u\mathbf{b}) = \mu, \tag{1b}$$

in the sense of distributions on $\mathbb{R}^d$. Characterize (compute) the distribution

$$\nu := \operatorname{div}(\beta(u)\mathbf{b}),$$

where $\beta \colon \mathbb{R} \to \mathbb{R}$ is a fixed C^1 function.
In the smooth setting one can use the standard chain rule formula to get

$$\begin{aligned} \nu = \operatorname{div}(\beta(u)\mathbf{b}) &= \beta'(u) \operatorname{div}(u\mathbf{b}) + (\beta(u) - u\beta'(u)) \operatorname{div} \mathbf{b} \\ &= \beta'(u)\mu + (\beta(u) - u\beta'(u))\lambda. \end{aligned} \tag{2}$$

The extension of (2) to a non-smooth setting is far from being trivial and this is exactly the aim of the chain rule problem.

As noted in [ADLM07], if one replaces “divergence” by “derivative”, the problem boils down to the one of writing a chain rule for weakly differentiable functions (a theme that has been investigated in several papers, see e.g. [Vol67, ADM90] for the BV setting). However, the “divergence” problem seems to be harder than the “derivative” one, due to stronger cancellation effects.
Problem 1 arises naturally in the study of partial differential equations, like the transport equation, the continuity equation or, more generally, hyperbolic conservations laws: indeed, they all can be written in the form $\operatorname{div}(uB) = c$, where $B \colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R} \times \mathbb{R}^d$ is vector field which has a space-time structure and $c \in \mathscr{D}'(\mathbb{R}^d)$ is some distribution. For instance, considering Problem 1 for a particular choice of B and β, one can establish uniqueness and comparison principles for weak solutions of scalar conservation laws (in the spirit of Kružkov’s theory, see [Kru70]).

1.1 Positive results

If we assume Sobolev regularity on the vector field, i.e. $\mathbf{b} \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^d)$ and $u \in L^q_{\mathrm{loc}}(\mathbb{R}^d)$ with p, q dual exponents, the chain rule has been established in [DL89]. In this case, it turns out that ν can be computed in terms of λ and μ just as in the classical (smooth) setting: it holds

$$\nu = (\beta(u) - u\beta'(u))\lambda + \beta'(u)\mu,$$

provided $\mu = \operatorname{div}(u\mathbf{b})$ is absolutely continuous with respect to the Lebesgue measure $\mathscr{L}^d$ on $\mathbb{R}^d$. This result has been extended in [ADLM07] to the case of a bounded variation vector field $\mathbf{b} \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$ and of a bounded density $u \in L^\infty_{\mathrm{loc}}(\mathbb{R}^d)$. More precisely, using the commutator estimate due to Ambrosio [Amb04], in [ADLM07] it is first proved that the distribution $\nu = \operatorname{div}(\beta(u)\mathbf{b})$ is a Radon measure which satisfies $\nu \ll |\lambda| + |\mu|$. Furthermore, the authors decompose λ, μ, ν into three parts (the *absolutely continuous* part λ^a, the *jump part* λ^j and the *Cantor part* λ^c, as in the standard BV setting) and treat them separately. They obtained that:

- the **absolutely continuous** part behaves as in the Sobolev case:

$$\nu^a = (\beta(u) - u\beta'(u))\lambda^a + \beta'(u)\mu^a, \qquad \text{as measures on } \mathbb{R}^d.$$

- For the **jump part**, they use the results obtained in [ACM05] to prove that ν^j can be computed in terms of the *traces* u^+ and u^- of u on the (countably) rectifiable set Σ where λ^j and μ^j are concentrated on.
- The **Cantor part** is harder and it is not characterized completely in [ADLM07], but only up to an error term. More precisely, they proved

$$\nu^c = (\beta(\widetilde{u}) - \widetilde{u}\beta'(\widetilde{u}))\lambda^c \llcorner_{\Omega \setminus S_u} + \beta'(\widetilde{u})\mu^c \llcorner_{\Omega \setminus S_u} + \sigma$$

  where $\widetilde{u}$ is the L^1 approximately continuous representative of u, S_u is the set of points where the L^1 approximate limit does not exist and σ is an error term (which is a measure concentrated on S_u, with $\sigma \ll \lambda^c + \mu^c$).

Further results in this directions have been obtained in [BG16], where the problem is completely solved in the case $d = 2$ with $\mathbf{b}$ of bounded variation, and in the recent preprint [BB17] where the analysis is completed (in the BV setting) for every $d \geq 2$.

1.2 Negative results

If we assume no regularity on $\mathbf{b}$ and u apart from measurability and boundedness, it can happen that λ and μ give no information about ν. This is related to the so-called problem of (non) *locality of the divergence operator*: indeed, in [ABC13] the authors constructed an example of a bounded vector field $\mathbf{v}$ (defined in $\mathbb{R}^2$) such that $\operatorname{div}\mathbf{v} \neq 0$, $\operatorname{div}\mathbf{v} \in L^\infty$ and $\operatorname{div}\mathbf{v}$ is supported on the set where $\mathbf{v}$ vanishes. Notice that this phaenomenon cannot occur for distributional derivatives, as they enjoy locality properties [AFP00, Prop. 3.73].
In the same spirit, in the recent work [CGSW17], using the abstract machinery of convex integration, the authors construct several examples of vector fields $\mathbf{b}\colon \mathbb{R}^d \to \mathbb{R}^d$ and densities $u\colon \mathbb{R}^d \to \mathbb{R}$ such that $\lambda = 0$, $\mu = 0$ but $\operatorname{div}(u^2\mathbf{b}) \neq 0$ in the sense of distributions in $\mathbb{R}^d$ for $d \geq 3$. More precisely, they show the following

Theorem 1 ([CGSW17]). *Let $d \geq 3$ and $\Omega \subset \mathbb{R}^d$ a smooth domain. Let f be a distribution such that the equation* $\operatorname{div} \mathbf{w} = f$ *admits a bounded, continuous solution* $\mathbf{w}\colon \Omega \to \mathbb{R}^d$ *on Ω. Then there exists a bounded vector field $\mathbf{b} \in L^\infty(\Omega;\mathbb{R}^d)$ and a density $u\colon \mathbb{R}^d \to \mathbb{R}$, with $0 < C^{-1} \leq u \leq C$ a.e. for some constant $C > 0$, such that*

$$\begin{aligned} &\operatorname{div} \mathbf{b} = 0 \\ &\operatorname{div}(u\mathbf{b}) = 0 \\ &\operatorname{div}\left(u^2\mathbf{b}\right) = f \end{aligned}$$

in the sense of distributions in Ω.

1.3 The two-dimensional case

The aim of this note is to address the two-dimensional case, i.e. $d = 2$. Notice that the assumption $d \geq 3$ is essential in [CGSW17], in view of the result of [BG16]. More precisely, in [BG16], the authors proved that if $d = 2$, $\mathbf{b}$ is bounded and of class BV and $u\colon \mathbb{R}^d \to \mathbb{R}$, with $0 < C^{-1} \leq u \leq C$ a.e. for some constant $C > 0$, are such that

$$\begin{aligned} &\operatorname{div} \mathbf{b} = 0 \\ &\operatorname{div}(u\mathbf{b}) = 0 \end{aligned}$$

then the Chain rule property holds, i.e. we have necessarily $\operatorname{div}\left(u^2\mathbf{b}\right) = 0$. Actually, the same conclusion is true if the assumption $\mathbf{b} \in \mathrm{BV}$ is replaced by $\mathbf{b} \neq 0$ a.e. in Ω. However, still remaining in the planar setting, in view of the results obtained in [BBG16], it seems reasonable to consider the chain rule also in the non steady setting, i.e. assuming that the vector field has a (special) space-time structure (and letting the divergence operator acting also on the time variable). More precisely, we are led to consider the following variant of Problem 1:

Problem 2. (Non steady chain rule) Let $T > 0$ be fixed and assume that it is given a bounded, Borel vector field $\mathbf{b}\colon \mathbb{R}^d \to \mathbb{R}^d$, a bounded, scalar function $u\colon (0,T) \times \mathbb{R}^d \to \mathbb{R}$ and Radon measures λ and μ such that

$$\operatorname{div} \mathbf{b} = \lambda,$$

$$\partial_t u + \operatorname{div}(u\mathbf{b}) = \mu,$$

in the sense of distributions on $(0,T) \times \mathbb{R}^d$. Characterize (compute) the distribution

$$\nu := \partial_t \beta(u) + \operatorname{div}(\beta(u)\mathbf{b}),$$

where $\beta : \mathbb{R} \to \mathbb{R}$ is a fixed C^1 function.

In this note we want to show the following

Theorem 2. *There exists an autonomous, compactly supported vector field* $\mathbf{b}\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{b} \in L^\infty(\mathbb{R}^2)$*, and a bounded, scalar function* $u\colon (0,T)\times\mathbb{R}^2 \to \mathbb{R}$*, such that*

$$\operatorname{div}\mathbf{b} = 0,$$

$$\partial_t u + \operatorname{div}(u\mathbf{b}) = 0,$$

in $\mathscr{D}'((0,T)\times\mathbb{R}^2)$ *but the distribution*

$$\partial_t\left(u^2\right) + \operatorname{div}\left(u^2\mathbf{b}\right) \notin \mathscr{M}((0,T)\times\mathbb{R}^2)$$

i.e. it is not *(representable by) a Radon measure.*

2 Preliminaries

In order to fix the notation, we collect in this section some preliminary results we will be using in the rest of the paper.

2.1 A particular change of variables

As in [ABC14, § 2.16], we will denote by I the interval $[0,L]$, by $\mathscr{L}^1$ the Lebesgue measure on I and, in general, λ will be an arbitrary measure on I, which is singular with respect to $\mathscr{L}^1$ and has $\mathscr{A}$ as the set of its atoms (points with positive measure). We set $\hat{L} := (\mathscr{L}^1 + \lambda)(I)$ and $\hat{I} := [0,\hat{L}]$. We denote by $\hat{\mathscr{L}}^1$ the Lebesgue measure restricted to $\hat{I}$. We denote by $\hat{\sigma}$ the multifunction from I to $\hat{I}$ that to every $s \in I$ associates the interval

$$\hat{\sigma}(s) := [\hat{\sigma}_-(s), \hat{\sigma}_+(s)]$$

where

$$\hat{\sigma}_-(s) := (\mathscr{L}+\lambda)\big([0,s)\big), \qquad \hat{\sigma}_+(s) := (\mathscr{L}+\lambda)\big([0,s]\big).$$

It is immediate to see that $\hat{\sigma}$ is surjective on I, strictly increasing, and uni-valued for every $s \notin \mathscr{A}$, because σ_- and σ_+ are strictly increasing, and $\sigma_-(s) = \sigma_+(s)$ whenever $s \notin \mathscr{A}$. Moreover it is obvious that the map is expanding, i.e.

$$s_2 - s_1 \le \hat{s}_2 - \hat{s}_1 \tag{5}$$

for every $s_1, s_2 \in I$ with $s_1 < s_2$, and every $\hat{s}_1 \in \hat{\sigma}(s_1), \hat{s}_2 \in \hat{\sigma}(s_2)$. Accordingly σ is surjective from $\hat{I}$ onto I, uni-valued and 1-Lipschitz (because of (5)); furthermore, it is constant on the interval $\sigma(s)$ for every $s \in \mathscr{A}$ and strictly increasing at every point outside $\sigma(\mathscr{A})$.

We recall the following

Lemma 1 ([ABC14, Lemma 2.17]). *Let F a $\mathscr{L}^1$-null set in I which supports the measure λ and $\hat{F} := \hat{\sigma}(F)$. Then*

1. *it holds $\sigma_\# \hat{\mathscr{L}}^1 = \mathscr{L}^1 + \lambda$;*
2. *the derivative of σ agrees with $\mathbf{1}_{\hat{I}\setminus\hat{F}}$ a.e. in $\hat{I}$.*

2.2 Solutions to singular, one-dimensional transport equations

In the following we will be dealing with 1d-transport equations involving singular terms, i.e. equations of the form

$$\partial_t \big(v(1 + \mathscr{L}^1 \times \lambda) \big) + \partial_s v = 0, \tag{6}$$

where $v\colon [0,T] \times I \to \mathbb{R}$ is a function of t,s and λ is a singular measure on I. Clearly, equation (6) has to be understood in the sense of distributions on $(0,T) \times I$: we say that v is a solution to (6) if for every $\phi \in C_c^\infty\big((0,T)\times I\big)$ it holds

$$\int_0^T \int_I v(t,s)\big(\phi_t(t,s) + \phi_s(t,s)\big)\, ds\, dt = -\int_0^T \int_I \phi_t(t,s)\, d\lambda(s)\, dt.$$

It is very well known that such equations present a severe phenomenon of non-uniqueness (for the associated initial value problem). In order to clarify what we mean, we begin by discussing an example.

Assume for simplicity that $I = \mathbb{R}$ and λ is the Dirac mass at 0, so that we are considering the equation

$$\partial_t \big(v(\mathscr{L}^1 \times \delta_0) \big) + \partial_s v = 0 \qquad \text{in } \mathscr{D}'((0,T)\times\mathbb{R}). \tag{7}$$

If v represents the density of a distribution of particles, then equation (7) is saying that each particle moves at constant speed 1 from left to right, except when it reaches the point 0, where it may stop for any given amount of time. Therefore, if v_0 is an arbitary, bounded initial datum (for simplicity, suppose its support is contained in $(-\infty,0)$), then a solution of (7) with initial condition $v(0,s) = v_0(s)$ is the function $v\colon [0,+\infty)\times\mathbb{R}\to\mathbb{R}$ defined by

$$v(t,s) = \begin{cases} v_0(s-t) & s \neq 0 \\ 0 & s = 0 \end{cases},$$

which physically means that no particle stops at 0. Another solution can be constructed by stopping all particles at 0, i.e.

$$\widetilde{v}(t,s) = \begin{cases} v_0(s-t) & s<0 \\ 0 & s=0 \\ \displaystyle\int_{-t}^{0} u_0(\tau)\,d\tau & s>0 \end{cases}.$$

More in general, for every $\alpha > 0$ one can construct a solution for which the particles arrive at 0, stay there exactly for time α and then leave (see Figure 1):

$$u^{\alpha}(t,s) := \begin{cases} u_0(s-t) & s<0 \\ \displaystyle\int_{-t}^{-t+\alpha} u_0(\tau)\,d\tau & s=0 \\ u_0(s-t+\alpha) & s>0. \end{cases}$$

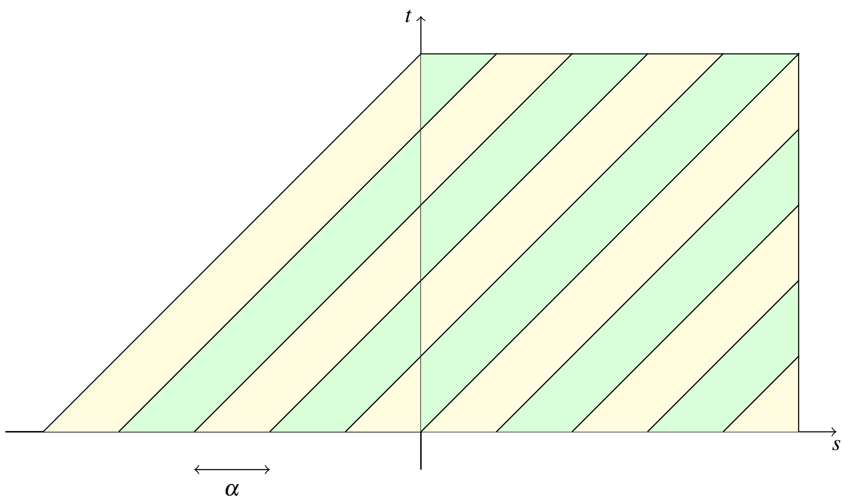


Fig. 1 A particular solution to equation (7): the particles at the initial time are of two different colors (yellow and green): they start moving following characteristic lines, arrive at 0 and stay there for a prescribed time α before leaving.

More precisely, we recall the following result, which is used in the proof of [ABC14, Lemma 4.5].

Lemma 2. *Let λ be a non trivial measure on $[0,L]$, singular w.r.t. to $\mathscr{L}^1 \llcorner_{[0,L]}$. Let furthermore $K \subset (0,L)$ be a closed, $\mathscr{L}^1$-negligible set, with $\lambda(K) > 0$. Then the problem*

$$\begin{cases} \partial_t\big(v(1+\mathscr{L}^1 \times \lambda)\big) + \partial_s v = 0 \\ v(0,\cdot) = \mathbf{1}_K(\cdot) \end{cases} \tag{8}$$

admits a non trivial bounded solution.

We recall here the main steps of the proof, as it will be useful in the following.

Proof. Clearly, the function $v(t,s) := \mathbf{1}_K(s)$ is a stationary solution of (8). Following [ABC14], we construct a second solution by exploiting the change of variable $s = \sigma(\hat{s})$ defined in 2.1. We thus define

$$v(t,s) := \begin{cases} w(t,\hat{\sigma}(s)) & \text{for } s \notin \mathscr{A} \\ \fint_{\hat{\sigma}(s)} w(t,\hat{s})\, d\hat{s}, & \text{for } s \in \mathscr{A}, \end{cases} \tag{9}$$

where we recall $\mathscr{A}$ is the set of atoms of λ and $w : [0,T) \times [0,L] \to \mathbb{R}$ is the (unique) bounded, distributional solution of

$$\begin{cases} \partial_t w + \partial_s w = 0 \\ w(0,\cdot) = \mathbf{1}_{\hat{\sigma}(K)}(\cdot). \end{cases} \tag{10}$$

To see that (9) actually solves Problem 8 we proceed as follows: first observe that (8) can be explicitly written as

$$\int_0^T \int_0^L \big(\partial_t \phi + \mathbf{1}_{I\setminus F} \partial_s \phi\big) v\, d(\mathscr{L}^1 + \lambda)\, dt = \int_0^L \phi(0,\cdot) \mathbf{1}_K d(\mathscr{L}^1 + \lambda). \tag{11}$$

By changing variable $s = \sigma(\hat{s})$, i.e. setting $\hat{v}(t,\hat{s}) := v(t,\sigma(\hat{s}))$ and $\hat{\phi}(t,\hat{s}) := \phi(t,\sigma(\hat{s}))$ and using Lemma 1, we can rewrite (11) as

$$\int_0^T \int_0^{\hat{L}} \big(\partial_t \hat{\phi} + \partial_{\hat{s}} \hat{\phi}\big) \hat{v}\, d\hat{s}\, dt = \int_0^{\hat{L}} \hat{\phi}(0,\cdot) \mathbf{1}_{\hat{\sigma}(K)} d\hat{s}.$$

Since on the complement of $\hat{\sigma}(\mathscr{A})$ it holds $\hat{v} = w$, to conclude we only need to show that

$$\int_{\hat{\sigma}(\mathscr{A})} \big(\partial_t \hat{\phi} + \partial_{\hat{s}} \hat{\phi}\big) \hat{v} d\hat{s} = \int_{\hat{\sigma}(J)} \big(\partial_t \hat{\phi} + \partial_{\hat{s}} \hat{\phi}\big) w d\hat{s}.$$

Indeed,

$$\begin{aligned} \int_{\hat{\sigma}(\mathscr{A})} (\partial_t \hat{\phi} + \partial_{\hat{s}} \hat{\phi}) \hat{v} d\hat{s} &= \sum_{a \in \mathscr{A}} \int_{\hat{\sigma}(a)} (\partial_t \hat{\phi} + \partial_{\hat{s}} \hat{\phi}) \hat{v} d\hat{s} \\ &= \sum_{a \in \mathscr{A}} \partial_t \phi(t,s) \int_{\hat{\sigma}(a)} \hat{v} d\hat{s} \\ &= \sum_{a \in \mathscr{A}} \partial_t \phi(t,s) \int_{\hat{\sigma}(a)} w = \int_{\hat{\sigma}(\mathscr{A})} (\partial_t \hat{\phi} + \partial_{\hat{s}} \hat{\phi}) w d\hat{s}, \end{aligned}$$

since $\partial_{\hat{s}} \hat{\phi}(t,\hat{s}) = 0$ and $\partial_t \hat{\phi}(t,\hat{s}) = \partial_t \phi(t,s)$ for all $\hat{s} \in \hat{\sigma}(s)$ and by direct definition of $\hat{v}$. To conclude the proof it is enough to show that the solution $\hat{v}$ does not coincide with the stationary one, and for this a possible strategy is to show that the maximum

$M(t)$ of the support of $v(t,\cdot)$ is strictly increasing at $t = 0$ (see [ABC14, Lemma 4.5])).

2.3 Structure of level sets of Lipschitz functions and Weak Sard Property

Since we will need some results on the structure of level sets of Lipschitz functions defined in the plane, we recall them here. Suppose that $\Omega \subset \mathbb{R}^2$ is an open, simply connected domain and $H\colon \Omega \to \mathbb{R}$ is a compactly supported Lipschitz function. For any $h \in \mathbb{R}$, let $E_h := H^{-1}(h)$. We recall the following deep

Theorem 3 ([ABC13, Theorem 2.5]). *Then the following statements hold for $\mathscr{L}^1$-a.e. $h \in H(\Omega)$:*

(1) $\mathscr{H}^1(E_h) < \infty$ *and E_h is countable $\mathscr{H}^1$-rectifiable (in what follows, we will say E_h is* regular*);*
(2) *for $\mathscr{H}^1$-a.e. $x \in E_h$ the function H is differentiable at x with* $\nabla H(x) \neq 0$*;*
(3) $\mathrm{Conn}^\star(E_h)$ *is countable and every $C \in$* $\mathrm{Conn}^\star(E_h)$ *is a closed simple curve;*
(4) $\mathscr{H}^1(E_h \setminus E_h^\star) = 0$.

Let us now define the critical set S of H as the set of all $x \in \Omega$ where H is not differentiable or $\nabla H(x) = 0$. We will be interested in the following *Weak Sard Property*, introduced in [ABC14, Section 2.13]:

$$H_\# \left(\mathscr{L}^2 \llcorner_{S \cap E^\star}\right) \perp \mathscr{L}^1,$$

where the set $E^\star$ is the union of all connected components with positive length of all level sets of H. The relevance of the Weak Sard Property in the framework of transport and continuity equation has been completely understood in [ABC14], to whom we refer the reader for further details. Here we simply point out that it is possible to prove that in some sense the Weak Sard Property is satisfied by a generic Lipschitz function (in Baire's category sense), as the class of all Lipschitz functions $H\colon \Omega \to \mathbb{R}$ satisfying the Weak Sard Property is residual in the Banach space of Lipschitz functions $\mathrm{Lip}(\Omega)$ (see [ABC13, Thm. 4]). However, an explicit construction of a Lipschitz function f without the Weak Sard property was proposed in [ABC13] and we will recall it in the Section 3, as it will be the building block of our counterexample.

2.4 Local disintegration of Lebesgue measure and of the equation $\mathrm{div}(u\mathbf{b}) = \mu$

From now onwards, let $\mathbf{b}\colon \Omega \to \mathbb{R}^2$ a bounded, Borel, divergence-free vector field. From $\mathrm{div}(\mathbf{b}) = 0$ in Ω we deduce that there exists a Lipschitz potential

$H\colon \Omega \to \mathbb{R}$ such that

$$\nabla^{\perp} H(x) = \mathbf{b}(x), \qquad \text{for } \mathscr{L}^2\text{-a.e. } x \in \Omega.$$

Using Theorem 3 on the Lipschitz function H, we can define the negligible set N_1 such that E_h is regular in Ω whenever $h \notin N_1$; moreover, let N_2 denote the negligible set on which the measure $(H_\# \mathscr{L}^2)^{\text{sing}}$ is concentrated, where $(H_\# \mathscr{L}^2)^{\text{sing}}$ is the singular part of $(H_\# \mathscr{L}^2)$ with respect to $\mathscr{L}^1$. Then we set

$$N := N_1 \cup N_2 \qquad \text{and} \qquad E^* := \cup_{h \notin N} E_h^* \tag{12}$$

For any $x \in E$ let C_x denote the connected component of E such that $x \in C_x$. By definition of E for any $x \in E$ the corresponding connected component C_x has strictly positive length. We recall the following Lemma which studies the disintegration of the measure $\mathscr{L}^2 \llcorner_{\Omega}$ w.r.t. the map H:

Lemma 3 ([ABC14, Lemma 2.8]). *There exist Borel families of measures σ_h, κ_h, $h \in \mathbb{R}$, such that*

$$\mathscr{L}^2 \llcorner_{\Omega} = \int \left(c_h \mathscr{H}^1 \llcorner_{E_h} + \sigma_h \right) dh + \int \kappa_h \, d\zeta(h), \tag{13}$$

where

1. $c_h \in L^1(\mathscr{H}^1 \llcorner_{E_h^\star})$, $c_h > 0$ *a.e.; moreover, by Coarea formula, we have* $c_h = 1/|\nabla H|$ *a.e. (w.r.t.* $\mathscr{H}^1 \llcorner_{E_h^\star}$*);*
2. σ_h *is concentrated on* $E_h^\star \cap \{\nabla H = 0\}$ *and* $\sigma_h \perp \mathscr{H}^1$ *for* $\mathscr{L}^1$*-a.e.* $h \notin N$*;*
3. κ_h *is concentrated on* $E_h^\star \cap \{\nabla H = 0\}$*;*
4. $\zeta := H_\# \mathscr{L}^2 \llcorner_{B \setminus E^\star}$ *is concentrated on* N *(hence* $\zeta \perp \mathscr{L}^1$*).*

2.4.1 Reduction of the equation on the level sets.

We now show how it is possible to reduce an equation of the form $\operatorname{div}(u\mathbf{b}) = \mu$, where u is a bounded Borel function on $\mathbb{R}^2$ and μ is a Radon measure on $\mathbb{R}^2$, into a family of 1d problems on the level sets of H. For all the following Lemmas we refer the reader to [BBG16].

The first step is the disintegration of the equation:

Lemma 4 ([BBG16, Lemma 3.5]). *Suppose that μ is a Radon measure on $\mathbb{R}^2$ and $u \in L^\infty(\mathbb{R}^2)$. Then the equation*

$$\operatorname{div}(u\mathbf{b}) = \mu \tag{14}$$

holds in $\mathscr{D}'(\Omega)$ if and only if:

- *the disintegration of μ with respect to H has the form*

$$\mu = \int \mu_h \, dh + \int \nu_h \, d\zeta(h), \tag{15}$$

where ζ is defined in Point (4) *of Lemma 3;*

- *for $\mathscr{L}^1$-a.e. $h \in \mathbb{R}$ it holds*

$$\operatorname{div}\left(u c_h \mathbf{b} \mathscr{H}^1 \llcorner_{E_h}\right) + \operatorname{div}(u \mathbf{b} \sigma_h) = \mu_h; \tag{16}$$

- *for ζ-a.e. $h \in \mathbb{R}$*

$$\operatorname{div}(u \mathbf{b} \kappa_h) = \nu_h. \tag{17}$$

2.4.2 Reduction on the connected components.

The next step is to reduce further the analysis of the equation (16) on the nontrivial connected components of the level sets. In view of Lemma 4 in what follows we always assume that $h \notin N$ (see (12)).

Lemma 5 ([BBG16, Lemma 3.7]). *The equation* (16) *holds iff*

- *for any nontrivial connected component C of E_h it holds*

$$\operatorname{div}\left(u c_h \mathbf{b} \mathscr{H}^1 \llcorner_C\right) + \operatorname{div}(u \mathbf{b} \sigma_{h \llcorner C}) = \mu_{h \llcorner C}; \tag{18}$$

- *it holds*

$$\operatorname{div}(u \mathbf{b} \sigma_{h \llcorner E_h \setminus E_h^\star}) = \mu_{h \llcorner E_h \setminus E_h^\star}. \tag{19}$$

Now we can split further and obtain the following

Lemma 6 ([BBG16, Lemma 3.8]). *Equation* (18) *holds iff*

$$\operatorname{div}\left(u c_h \mathbf{b} \mathscr{H}^1 \llcorner_C\right) = \mu_{h \llcorner C}, \tag{20a}$$

$$\operatorname{div}(u \mathbf{b} \sigma_{h \llcorner C}) = 0. \tag{20b}$$

2.4.3 Reduction of the equation on connected components in parametric form

Finally, we would like to discuss the parametric version of the equation (20a). Let $\gamma \colon I \to \mathbb{R}^2$ be an injective Lipschitz parametrization of C, where $I = \mathbb{R}/\ell\mathbb{Z}$ or $I = (0, \ell)$ (for some $\ell > 0$) is the domain of γ. The existence of such a parameterization is granted by [ABC13, Thm. 2.5 (iv)].

Lemma 7 ([BBG16, Lemma 3.9]). *Equation* (20a) *holds iff for any admissible parametrization γ of C*

$$\partial_s \hat{u}_h = \hat{\mu}_h \tag{21}$$

where $\gamma_\# \hat{\mu}_h = \mu_{h \llcorner C}$, $\hat{u}_h = u \circ \gamma$.

2.5 Local disintegration of a balance law

We now pass to consider a general balance law associated to the Hamiltonian vector field $\mathbf{b}$, i.e. $\partial_t u + \mathrm{div}(u\mathbf{b}) = \nu$, being ν a Radon measure on $(0,T)\times\Omega$ and $u \in L^\infty((0,T)\times\Omega)$. A reduction on the connected components of the Hamiltonian H can be performed, similarly to what we have done for equation $\mathrm{div}(u\mathbf{b}) = \mu$ to above. In some sense, we are presenting now the time-dependent version of Lemmas 4-5-6-7.

Lemma 8. *A function $u \in L^\infty([0,T]\times\Omega)$ is a solution to the problem*

$$\begin{cases} u_t + \mathrm{div}(u\mathbf{b}) = \nu, \\ u(0,\cdot) = u_0(\cdot), \end{cases} \qquad \text{in } \mathscr{D}'((0,T)\times\Omega) \tag{22}$$

if and only if

- $\hat{u}_h(t,s) := u(t,\gamma_h(s))$ *solves*

$$\begin{cases} \partial_t \hat{u}_h + \partial_s \hat{u}_h = \hat{\nu}_h \\ \hat{u}_h(0,\cdot) = \hat{u_0}_h(\cdot), \end{cases} \qquad \text{in } \mathscr{D}'((0,T)\times I)$$

- *it holds*

$$\mathrm{div}(u\mathbf{b}\sigma_h) = 0$$

for $\mathscr{L}^1$-a.e. h, where $\gamma_h \colon I \to \mathbb{R}^2$ is an admissible parametrization of a connected component C of the level set E_h of the Hamiltonian H and $\hat{\nu}_h$ is a measure such that $\hat{\nu}_h = (\gamma_h^{-1})_\# \nu$.

Proof. Multiplying equation in (22) by a function $\psi \in C_c^\infty([0,T))$ and formally integrating by parts we get

$$u_t\psi + \mathrm{div}(u\psi\mathbf{b}) = \psi\nu \Rightarrow \mathrm{div}\left(\int_0^T u\psi\,dt\,\mathbf{b}\right) = \int_0^T u\psi_t\,dt - \psi(0)u_0 + \left(\int_0^T \psi\,dt\right)\nu,$$

which can be written in the form

$$\mathrm{div}(w\mathbf{b}) = \mu, \tag{23}$$

where $w := \int_0^T u\psi\,dt$ and

$$\mu := \left(\int_0^T u\psi_t\,dt - \psi(0)v_0\right)\mathscr{L}^2 + \left(\int_0^T \psi\,dt\right)\nu.$$

Applying Lemma 4 and Lemma 6 to (23), we obtain that continuity equation is equivalent to

$$\mathrm{div}\left(wc_h\mathbf{b}\mathscr{H}^1 \llcorner E_h\right) = \mu_h \tag{24}$$

and

$$\operatorname{div}(u\mathbf{b}\sigma_h) = 0 \tag{25}$$

for $\mathscr{L}^1$-a.e. h, where the measure μ_h can be computed explicitly, using Coarea Formula and disintegration Theorem

$$\mu_h = \left(\int_0^T u\psi_t\,dt - \psi(0)v_0\right)\mathscr{H}^1 \llcorner E_h + \left(\int_0^T \psi\,dt\right)\nu_h.$$

Thanks to Lemma 7, equation (24) is *equivalent* to

$$\partial_s \hat{u} = \hat{\mu}_h,$$

in $\mathscr{D}'(I)$. Now being γ_h Lipschitz and injective, we have

$$(\gamma_h^{-1})_\#\big(\mathscr{H}^1 \llcorner E_h\big) = |\gamma_h'|\mathscr{L}^1,$$

and this allows us to compute explicitly

$$\begin{aligned}
\hat{\mu}_h &= (\gamma_h^{-1})_\#\mu_h \\
&= (\gamma_h^{-1})_\#\left(\int_0^T u\psi_t\,dt\,c_h\mathscr{H}^1 \llcorner E_h - \int_{\mathbb{R}^2}\psi(0)v_0c_hd\mathscr{H}^1 \llcorner E_h + \int_0^T \psi\,dt\,\nu_h\right) \\
&= \int_0^T v(\tau,\gamma(s))\psi_\tau(\tau)\,d\tau - \psi(0)u_0(\gamma_h(s))c_h(\gamma(s)) + \left(\int_0^T \psi(\tau)d\tau\right)\hat{\nu}_h,
\end{aligned} \tag{26}$$

where

$$\hat{\nu}_h = (\gamma_h^{-1})_\#\nu.$$

Formally, (26) means

$$\hat{\mu}_h = -\int_0^T \partial_t\hat{u} + \hat{\nu}_h.$$

To sum up, we have obtained that Problem (22) is equivalent to

$$\begin{cases}\partial_t\hat{u}_h + \partial_s\hat{u}_h = \hat{\nu}_h,\\ \hat{u}_h(0,\cdot) = \hat{u_0}_h(\cdot),\end{cases}$$

and

$$\operatorname{div}(u\mathbf{b}\sigma_h) = 0$$

in $\mathscr{D}'((0,T)\times I)$ for $\mathscr{L}^1$-a.e. $h\in\mathbb{R}$. We explicitly notice that the last relation is always satisfied, as $\mathbf{b}=0$ on the critical set (where σ_h is concentrated).

2.6 Anzellotti traces for measure divergence L^∞ vector fields

Let now $\rho \in L^\infty(\mathbb{R}^d)$ and $\mathbf{v} \in L^\infty(\mathbb{R}^d;\mathbb{R}^d)$. We consider the vector field $\mathbf{V} := \rho\mathbf{v}$ and we assume that its distributional divergence $\operatorname{div}\mathbf{V}$ is represented by some Radon measure, so that $\mathbf{V}$ is a so called *measure divergence vector field*. There are well known results that allows to give a meaning and to characterize the *trace* of such vector fields over rectifiable sets. We list here the main ones and we refer for more details to [Anz83, DL07, ACM05].

Definition 1. Given a bounded, open domain with C^1 boundary $U \subset \mathbb{R}^d$, the *(Anzellotti) normal trace* of $\mathbf{V}$ over ∂U is the distribution defined by

$$\big\langle \mathsf{Tr}(\mathbf{V},U)\cdot\mathbf{n},\psi\big\rangle := \int_U \psi(x)\,d(\operatorname{div}\mathbf{V})(x) + \int_U \mathbf{V}\cdot\nabla\psi(x)\,d\mathscr{L}^d(x)$$

for every compactly supported smooth test function $\psi \in C_c^\infty(\mathbb{R}^d)$.

We have the following Proposition, which says that the trace of a measure divergence vector field is not an arbitray distribution, but is is induced by integration of a bounded function defined on ∂U.

Proposition 1. *There exists a unique $g \in L^\infty_{\mathrm{loc}}(\partial U;\mathscr{H}^{d-1}\llcorner_{\partial U})$ such that*

$$\big\langle \mathsf{Tr}(\mathbf{V},U)\cdot\mathbf{n},\phi\big\rangle = \int_{\partial U} g\phi\,\mathscr{H}^{d-1}, \qquad \forall\phi \in C_c^\infty(\mathbb{R}^d).$$

One can also define the traces of $\mathbf{V}$ on a *oriented* hypersurface of class C^1, say Σ. Indeed, choosing an open C^1 domain $U \Subset U$ such that $\Sigma \subset \partial U$ and the unit outer normals agree $\nu_U = \nu_\Sigma$ we can define

$$\mathsf{Tr}^-(\mathbf{V},\Sigma)\cdot\mathbf{n} := \mathsf{Tr}(\mathbf{V},U).$$

Analogously, choosing an open C^1 domain U' such that $\Sigma \subset \partial U'$ and $\nu_{U'} = -\nu_\Sigma$ we define

$$\mathsf{Tr}^+(\mathbf{V},\Sigma)\cdot\mathbf{n} := -\mathsf{Tr}(\mathbf{V},U')\cdot\mathbf{n}.$$

We remark that one can replace C^1 regularity with Lipschitz, so that it is possible to give the definition of normal trace of a measure divergence vector field on countable H^{d-1}-rectifiable sets.

We collect here other important results on Anzellotti's weak traces:

Proposition 2. *If $\mathbf{V}$ is a bounded, measure divergence vector field, then:*

- $\operatorname{div}\mathbf{V} \ll \mathscr{H}^{d-1}$ *as measures in $\mathbb{R}^d$;*
- *for any oriented, C^1 hypersurface Σ it holds*

$$\operatorname{div}\mathbf{V}\llcorner_\Sigma = \big(\mathsf{Tr}^+(\mathbf{V},\Sigma)\cdot\mathbf{n} - \mathsf{Tr}^-((\mathbf{V},\Sigma)\cdot\mathbf{n}\big)\mathscr{H}^{d-1}\llcorner_\Sigma.$$

Finally, an interesting case is when we assume more regularity on the vector field, for instance $\mathbf{v} \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$. In this situation, one has the usual definition of the trace of $\mathbf{v}$ over ∂U as BV function. We recall that the trace of BV functions $\mathbf{v}$ for open sets $U \subset \mathbb{R}^d$ of class C^1 is a measure which is absolutely continuous w.r.t. $\mathscr{H}^{d-1} \llcorner_{\partial U}$. We conclude this section by recalling the following chain rule for traces, proved when $\mathbf{v} \in \mathrm{BV}$ in [ADLM07] (see also [ACM05] for the case of vector fields of *bounded deformation*).

Theorem 4 (Change of variables for traces). *Let $U \subset \mathbb{R}^d$ be an open domain of class C^1 and let $\mathbf{v} \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$ and $\beta \in \mathrm{Lip}(\mathbb{R})$. Then if $\mathbf{V} = \rho\mathbf{v}$ is a measure divergence vector field, then also $\beta(\rho)\mathbf{v}$ is a measure divergence vector field and, moreover, it holds*

$$\mathsf{Tr}^{\pm}(\beta(\rho)\mathbf{v},U)\cdot\mathbf{n} = \beta\left(\frac{\mathsf{Tr}^{\pm}(\rho\mathbf{v},U)\cdot\mathbf{n}}{\mathsf{Tr}^{\pm}(\mathbf{v},U)\cdot\mathbf{n}}\right)\mathsf{Tr}^{\pm}(\mathbf{v},U)\cdot\mathbf{n}, \qquad \mathscr{H}^{d-1}\text{-a.e. on } \partial U,$$

where the ratio is arbitrarily defined at points where the trace $\mathsf{Tr}(\mathbf{v},U)$ vanishes.

3 The construction of the Hamiltonian

In this section, we construct a suitable Lipschitz function $H\colon \mathbb{R}^2 \to \mathbb{R}$ that will be the building block of our counterexample. The construction presented in this paragraph goes back to [ABC13].

3.1 A function that does not have Weak Sard Property

Let $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ be decreasing sequences of positive numbers with asymptotic behaviour given by

$$a_n \sim b_n \sim \frac{1}{n^2 2^n}.$$

Hence, the following quantities

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

are finite. Chosen a real number $\delta > 0$, we set

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

3.1.1 The construction of the set

We consider the set C_0, which is the closed rectangle with width c_0 and height d_0. Then we define C_1 to be the union of 4 closed rectangles with sizes

$$c_1 := \frac{c_0}{2} - 2a_0, \qquad d_1 := \frac{d_0}{2} - b_0$$

like in Figure 2.

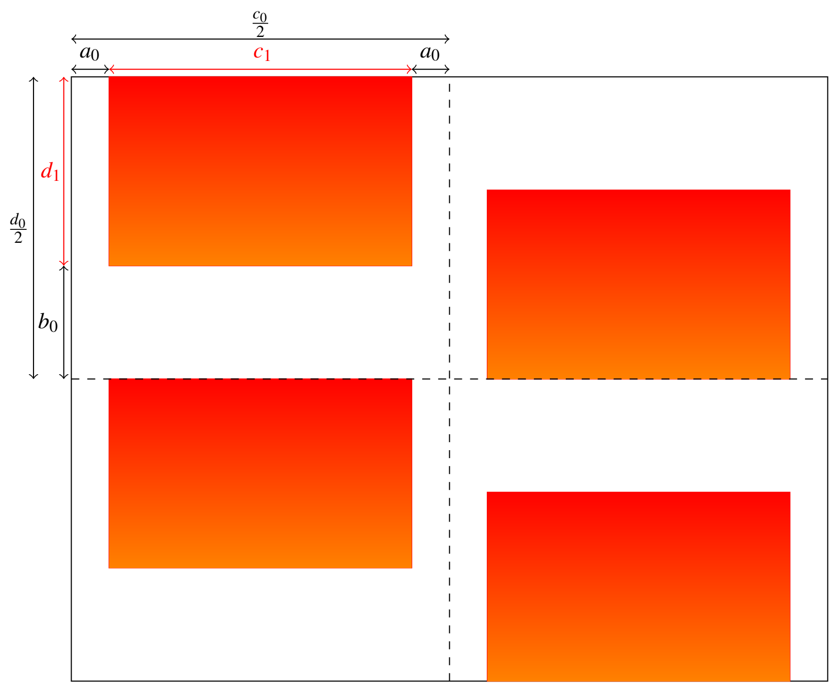


Fig. 2 The sets C_0 (black) and C_1 (red).

If we iterate the above construction, we obtain a sequence of nested sets: more precisely, if C_n is the union of 4^n pairwise disjoint, closed rectangles with width c_n and height d_n, then C_{n+1} is the union of 4^{n+1} pairwise disjoint closed rectangles with width

$$c_{n+1} := \frac{c_n}{2} - 2a_n, \qquad d_{n+1} := \frac{d_n}{2} - b_n.$$

It is easy to see that from this recursion we have

$$2^n c_n = c_0 - \sum_{m=0}^{n-1} 2^{m+2} a_m \searrow \delta \qquad \text{and} \qquad 2^n d_n = d_0 - \sum_{m=0}^{n-1} 2^{m+2} b_m \searrow \delta$$

which implies that c_n, d_n are always strictly positive and satisfy

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

If C denotes the intersection of the closed sets C_n we have

$$\mathscr{L}^2(C) = \lim_n \mathscr{L}^2(C_n) = \lim_n 4^n d_n c_n = \delta^2.$$

3.1.2 Construction of the function

We now turn to the construction of a suitable sequence of Lipschitz and piecewise smooth functions $f_n \colon \mathbb{R}^2 \to \mathbb{R}$. The function f_0 is defined by its level sets, drawn in Figure 3a.

Let s_n be the oscillation of the function f_n on the component of C_n; it is clear from the picture that

$$s_{n+1} = \frac{s_n}{4}, \tag{27}$$

hence $s_n = 4^{-n} s_0 = 4^{-n} d_0$.

3.1.3 L^∞ gradient estimates

We can now estimate the gradient of the functions f_n. It is easy to see that the supremum of $|\nabla f_n|$ in the set C_n is attained in the set E defined in Figure 3a. Choosing the axes as in Figure 3b we can write an explicit formula for f_n; in particular, the line that passes through the points $(-a_n, b_n)$ and $(a_n, \frac{d_n}{2} - b_n)$ has equation

$$x_2 = b_n + \frac{1}{4a_n}(x_1 + a_n)(d_n - 4b_n).$$

Then if we pick a $\tau \in (0, b_n)$ we impose the similarity of the triangles, hence

$$\frac{\tau b_n}{b_n} = \frac{x_2}{b_n + \frac{1}{4a_n}(x_1 + a_n)(d_n - 4b_n)}$$

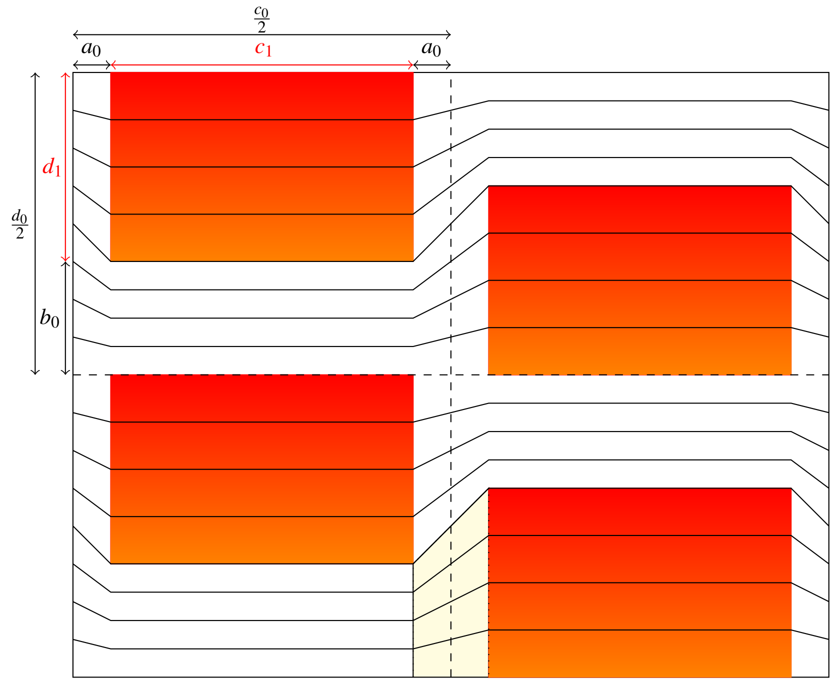
hence we get

$$\tau = \frac{4a_n x_2}{(d_n - 4b_n)x_1 + a_n d_n}.$$

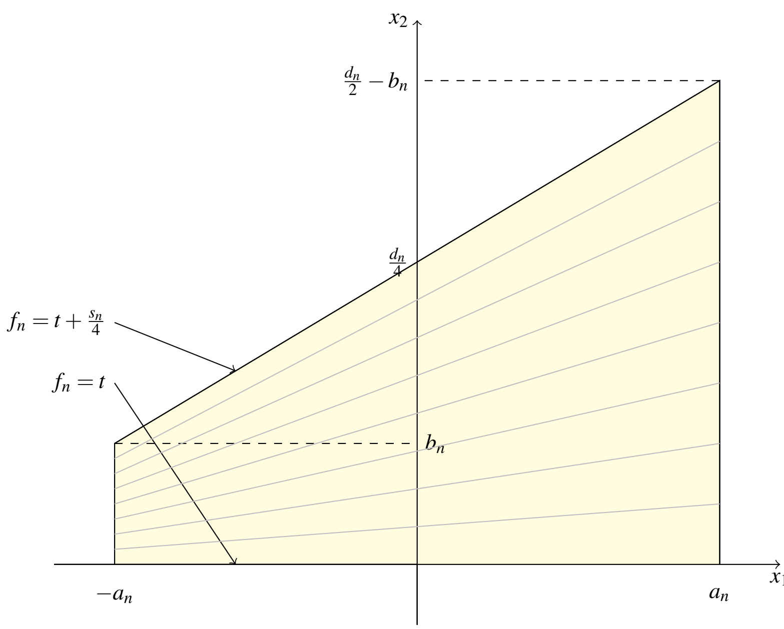
Therefore, the function f_n has the following explicit formula in E:

$$f_n(x_1, x_2) = (1-\tau)t + \tau\left(t + \frac{s_n}{4}\right) = t + \frac{s_n}{4}\tau = t + \frac{a_n s_n x_2}{(d_n - 4b_n)x_1 + a_n d_n}.$$

A direct computation shows that



(a) Level sets of the function f_0.



(b) Estimate of $|\nabla f_n|$: the level sets of f_n in the set E.

Fig. 3 Level sets of the function f_0 and estimates for $|\nabla f_n|$.

$$\nabla f_n(x) = \frac{1}{a_n d_n + (d_n - 4b_n)x_1} \left(-(d_n - 4b_n)(f_n(x) - t), a_n s_n\right).$$

Taking into account that $x_1 \geq -a_n$ and that $d_n - 4b_n > 0$ (due to the asymptotic behaviour) we can estimate from below the denominator:

$$a_n d_n + (d_n - 4b_n)x_1 \geq 4a_n b_n.$$

On the other hand, we clearly have $|f_n - t| \leq s_n$ and thus we obtain the following estimate:

$$\|\nabla f_n\|_{L^\infty(C_n)} \leq \frac{(d_n - 4b_n)s_n + a_n s_n}{4a_n b_n} = O(n^4 2^{-n}). \tag{28}$$

Now let us define the function $h_n := f_n - f_{n-1}$. Clearly, by definition of f_n, the support of h_n lies in C_n; moreover,

$$\|\nabla h_n\|_\infty \leq \|\nabla f_n\|_{L^\infty(C_n)} + \|\nabla f_{n-1}\|_{L^\infty(C_{n-1})} \sim n^4 2^{-n}. \tag{29}$$

Since the distance of a point in C_n from $\mathbb{R}^2 \setminus C_n$ is of order $c_n \sim 2^{-n}$, by the Mean Value Theorem

$$\|h_n\|_{L^\infty} \sim n^4 4^{-n}.$$

For every $x \in \mathbb{R}^2$ set

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

We sum up the properties of the function f in the following

Theorem 5 ([ABC13, Prop. 4.7]). *If C is the set above and f is the function built in the previous sections, then:*

(i) f is differentiable at every $x \in C$ with $\nabla f(x) = 0$;
(ii) $\mathscr{L}^1(f(C)) = d_0$;
(iii) $f_\sharp(\mathscr{L}^2 \llcorner C) = m\mathscr{L}^1 \llcorner f(C)$, where $m = \delta^2/d_0$; in particular, f does not satisfy the Weak Sard Property.

Proof. (i) The estimates (28) and (29) yield

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

which means that the Lipschitz constant of f on each component of C_n is of order $\mathscr{O}(n^4 2^{-n})$; being C contained in the interior of C_n, it follows that for every $x \in C$

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

Letting $n \to +\infty$ we obtain the assertion, i.e. f is differentiable at x with $\nabla f(x) = 0$.

(ii) The range $f(C)$ is the intersection of all $f(C_n)$, and $f(C_n)$ agrees with $f_n(C_n)$, and therefore it is the union of 4^n pairwise disjoint, closed intervals each of length s_n (defined as the oscillation of f_n in C_n). Thus from (27),

$$\mathscr{L}^1(f(C)) = \lim_n \mathscr{L}^1(f_n(C_n)) = \lim_n 4^n s_n = d_0.$$

(iii) We must show that the measures $\mu := f_\#(\mathscr{L}^2 \llcorner S)$ and $\lambda := m\mathscr{L}^1 \llcorner f(C)$ are the same. Since both μ and λ are supported on the compact set $f(C)$, we apply Lemma 4.6 in [ABC13] to the partitions F_n given by the sets $R' := f(R \cap C)$ where R is a component of C_n, and deduce that it suffices to prove $\mu(R') = \lambda(R')$ for every such R'. Since C can be written as a disjoint union of 4^n translated copies of $R \cap C$, we have

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

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

$$\lambda(R') = m\mathscr{L}^1(R') = 4^{-n}m\mathscr{L}^1(f(C)) = 4^{-n}md_0 = 4^{-n}\delta^2.$$

3.1.4 Further remarks on the Hamiltonian without Weak Sard Property

The Lipschitz function constructed in 3.1 will be denoted as $f_{c_0,d_0,\delta}$, since c_0, d_0, δ are free parameters in the construction. Recall also that $\mathrm{osc} f_{c_0,d_0,\delta} = d_0$ so that, up to a translation, we can suppose directly that

$$f_{c_0,d_0,\delta}(\mathbb{R}^2) = (0, d_0).$$

The critical set S of $f_{c_0,d_0,\delta}$ has area $\mathscr{L}^2(S) = \delta^2$ and, as shown in Theorem 5

$$(f_{c_0,d_0,\delta})_\#(\mathscr{L}^2 \llcorner C) = \frac{\delta^2}{d_0}\mathscr{L}^1 \llcorner f(C).$$

Therefore, we can apply Disintegration Theorem to the probability measure $\frac{1}{\delta^2}\mathscr{L}^2 \llcorner C$ w.r.t. the map $f_{c_0,d_0,\delta}$. We thus write

$$\frac{1}{\delta^2}\mathscr{L}^2 \llcorner C = \frac{1}{d_0}\int \nu_h \, dh$$

where $h \mapsto \nu_h$ is a measurable measure-valued map, ν_h being a probability measure concentrated on $f^{-1}_{c_0,d_0,\delta}(h) \cap C$ for $\mathscr{L}^1$-a.e. $h \in \mathbb{R}$. We can actually say more, characterizing completely the measure ν_h. In particular, we want to show that for a.e. h the intersection

$$f^{-1}_{c_0,d_0,\delta}(h) \cap C$$

is a single point. We have indeed

$$f^{-1}_{c_0,d_0,\delta}(h)\cap C=\bigcap_n\big(f^{-1}_{c_0,d_0,\delta}(h)\cap C_n\big)$$

and for every h it is possible to prove that $f^{-1}_{c_0,d_0,\delta}(h)\cap C_n$ is a sequence of nested intervals whose measure goes to 0 as $n\to+\infty$. For instance, if $h\in(d_0/2^n,d_0)$, we have that

$$f^{-1}(h)\cap C_n=\left(a_{n-1},\frac{c_{n-1}}{2}-a_{n-1}\right)\times\{d_0\}.$$

The lenght of the interval is clearly $\frac{c_{n-1}}{2}-2a_{n-1}=c_n\simeq\delta\cdot 2^{-n}\to 0$ as $n\to+\infty$. This shows that $f^{-1}(h)\cap C=\{x_h\}$ for every $h\in(0,d_0)=f([0,c_0]\times[0,d_0])$. So ν_h has to be δ_{x_h}. Finally notice that we can write

$$\mathscr{L}^2\llcorner_C=\int\delta_{x_h}m\,dh.$$

3.1.5 Scaling the Hamiltonian $f_{1,1,\delta}$

Set now $H_1:=f_{1,1,\delta}$ whose range is $(0,1)$. The disintegration now looks like

$$\mathscr{L}^2\llcorner_C=\delta^2\int\delta_{x_h}\,dh=|S|\int\delta_{x_h}\,dh$$

which will be written from now onwards as

$$\mathscr{L}^2\llcorner_C=\int c_h\delta_{x_h}\,dh,$$

where we have set for $\mathscr{L}^1$-a.e. h the coefficient $c_h:=|S|$. The map $h\mapsto c_h$ is thus constant and it simply represents the density of $f_\sharp(\mathscr{L}^2\llcorner_C)$ along the level sets. We will see that this map plays a significant role in the construction: we will suitably modify it, in order to obtain a piecewise constant map which is integrable but not square-integrable. To do this, we perform some scaling transformations: for fixed $n\in\mathbb{N}$ and $\alpha\in\mathbb{R}$ we first scale the *domain* of H_1 with the following linear map:

$$\mathscr{Q}_n\colon (x,y)\mapsto\left(x,\frac{y}{2^n}\right)$$

The area of the critical set was $|S|=\delta^2=\int_0^1 c_h\,dh$, while after the operation the area becomes

$$\det\mathscr{Q}_n\cdot|S|=\frac{|S|}{2^n}$$

hence we set

$$c'_h:=\frac{|S|}{2^n}.$$

Now we rescale the range $(0,1)\mapsto(0,2^{-n\alpha})$ via a map $\mathscr{R}_{n,\alpha}\colon\mathbb{R}\to\mathbb{R}$ so that if we impose

$$\frac{|S|}{2^n} = \int_0^{2^{-n\alpha}} c''_h \, dh$$

we have to set accordingly

$$c''_h := \frac{|S|}{2^n} 2^{n\alpha} = \frac{|S|}{2^{n(1-\alpha)}}$$

Finally, we define the translation operator $\mathscr{T}_{n,\alpha}$ which acts both in the domain and in the target in the following way: if a function is defined in the square $[0,1] \times \left[0, \frac{1}{2^n}\right] \subset \mathbb{R}^2$ with range $\left[0, \frac{1}{2^{n\alpha}}\right]$ then under the action of $\mathscr{T}_{n,\alpha}$ the domain becomes the rectangle $[0,1] \times \left[\frac{1}{2^n}, \frac{1}{2^{n-1}}\right]$ while the range turns to the interval $\left[\frac{1}{2^{n\alpha}}, \frac{1}{2^{n\alpha-1}}\right]$. We call the resulting function $\mathscr{T}_{n,\alpha} \circ \mathscr{R}_{n,\alpha} \circ \mathscr{Q}_{n,\alpha} \circ H_1 := H_{n,\alpha}$ and we define now

$$H_\alpha(x,y) := \sum_{n\in\mathbb{N}} H_{n,\alpha}(x,y), \qquad (x,y) \in D := \bigcup_{n\in\mathbb{N}} \left([0,1] \times \left[\frac{1}{2^n}, \frac{1}{2^{n-1}}\right] \right) = [0,1] \times [0,1].$$

In other words, we have "patched together" the rescaled Hamiltonians, one above the other, with ranges that are adiacent intervals. Notice that the function is well defined, as the domains of the different $H_{n,\alpha}$ are disjoint, so that for any $(x,y) \in D$ the sum is locally finite (actually it reduces to a single term).

3.1.6 Properties of H_α

Some remarks about the properties of H_α are now in order.

- For $\alpha > 0$, the function H_α is bounded. Indeed, its range is

$$H_\alpha(D) = \bigcup_{n\in\mathbb{N}} \left[2^{-n\alpha}, 2^{-n\alpha+1}\right]$$

whose measure is

$$\mathscr{L}^2(H_\alpha(D)) = \sum_n \frac{1}{2^{n\alpha}} =: \ell_\alpha < +\infty,$$

for $\alpha > 0$.

- For any $\alpha \in \mathbb{R}$, the area of the critical set of H_α is always finite:

$$\int_0^{\ell_\alpha} c''_h \, dh = \sum_n \frac{|S|}{2^{n(1-\alpha)}} \times \frac{1}{2^{n\alpha}} = \sum_n \frac{|S|}{2^n} = |S| < +\infty.$$

- On the contrary, we have that

$$\int_0^{\ell_\alpha} (c''_h)^2 \, dh = \sum_n \frac{|S|^2}{2^{2n(1-\alpha)}} \times \frac{1}{2^{n\alpha}} = \sum_n \frac{|S|^2}{2^{n(2-\alpha)}}.$$

In particular, if we take $\alpha \geq 2$ we have that

$$\int_0^{\ell_\alpha} (c_h'')^2 \, dh = +\infty.$$

In other words, for $\alpha \geq 2$, the function $h \mapsto c_h''$ belongs to $L^1([0,\ell_\alpha]) \setminus L^2([0,\ell_\alpha])$ (it behaves essentially like $n\mathbf{1}_{[0,n^{-2}]}$ in $[0,1]$).

4 The counterexample

We now fix $\alpha > 2$ and we consider the corresponding Hamiltonian H_α constructed in paragraph above and we set $\mathbf{b} := \nabla H_\alpha$. By construction, setting $\sigma_h := c_h'' \delta_{x_h}$, we have that H_α satisfies the following

$$\mathscr{L}^2 = \int_{\mathbb{R}} \left(\frac{1}{|\nabla H|} \mathscr{H}^1 \llcorner E_h + \sigma_h \right) dh. \tag{30}$$

For typographical reasons, we will write from now onward simply c_h instead of c_h''.

By applying Lemma 8 to H_α we get at once the following

Proposition 3. *The problem*

$$\begin{cases} \partial_t u + \operatorname{div}(u\mathbf{b}) = 0 \\ u(0,\cdot) = u_0(\cdot) \end{cases} \tag{31}$$

is equivalent to

$$\begin{cases} \partial_t \hat{u}_h + \partial_s \hat{u}_h + c_h \partial_t \hat{u} \mathscr{L}^1 \otimes \delta_{s_h} = 0 \\ \hat{u}_0(s) = u_{0h} \\ \partial_s \big(\hat{u}_h c_h \mathscr{L}^1 \otimes \delta_{s_h} \big) = 0 \end{cases} \tag{32}$$

for $\mathscr{L}^1$-a.e. h.

Remark 1. Notice that, by splitting

$$u = m\mathbf{1}_S + u\mathbf{1}_{S^c}$$

the equation can be written as

$$\partial_t (u\mathbf{1}_{S^c}) + \partial_t (m\mathbf{1}_S) + \underbrace{\operatorname{div}(m\mathbf{1}_S\mathbf{b})}_{=0} + \operatorname{div}(u\mathbf{1}_{S^c}\mathbf{b}) = 0$$

because $\mathbf{b} = 0$ on S by construction. Hence, taking into account that $\mathbf{b} = 0$ on the critical set, Proposition 3 is actually establishing that

$$\begin{cases} \partial_t(u\mathbf{1}_{S^c}) + \mathrm{div}(u\mathbf{1}_{S^c}\mathbf{b}) = -\partial_t(m\mathbf{1}_S) \\ u(0,\cdot) = u_0(\cdot) \end{cases}$$

is *equivalent* to

$$\begin{cases} \partial_t \hat{u}_h + \partial_s \hat{u}_h + c_h \partial_t \hat{m}_h \mathscr{L}^1 \otimes \delta_{s_h} = 0 \\ \hat{u}_0(s) = u_{0h} \end{cases} \qquad \text{for } \mathscr{L}^1\text{-a.e. } h.$$

Now we consider the Cauchy problem for the transport equation associated to $\mathbf{b}$ with initial condition $u_0 := \mathbf{1}_S$:

$$\begin{cases} \partial_t u + \mathrm{div}(u\mathbf{b}) = 0 \\ u(0,\cdot) = \mathbf{1}_S(\cdot) \end{cases} .$$

We disintegrate the equation on the level sets and we obtain, denoting for typographic simplicity by $v_h(t,s) := \hat{u}_h(t,s)$, we have

$$\begin{cases} \partial_t v_h + \partial_s v_h = -c_h \partial_t (v_h \mathscr{L}^1 \times \delta_{s_h}) \\ v_h(0,\cdot) = c_h \mathbf{1}_{\{s_h\}}(\cdot) \end{cases} .$$

i.e.

$$\begin{cases} \partial_t \big(v_h (1 + \mathscr{L}^1 \times c_h \delta_{s_h}) \big) + \partial_s v_h = 0 \\ v_h(0,\cdot) = c_h \mathbf{1}_{\{s_h\}}(\cdot) \end{cases} , \tag{33}$$

which is exactly of the form 8. Applying Lemma 2, we have that the function

$$v_h(t,s) := \begin{cases} c_h \mathbf{1}_{\hat{\sigma}(s_h)}(\hat{\sigma}(s) - t) & s \neq s_h \\ \displaystyle\fint_{\hat{\sigma}(s_h)} c_h \mathbf{1}_{\hat{\sigma}(s_h)}(\hat{s} - t) d\hat{s} & s = s_h \end{cases}$$

is a non-stationary solution to (33). Some easy computations show that

$$\begin{aligned} v_h(t,s_h) = \fint_{\hat{\sigma}(s_h)} \mathbf{1}_{\hat{\sigma}(s_h)}(\hat{s} - t) d\hat{s} &= \frac{1}{c_h} \int_{s_h}^{s_h + c_h} c_h \mathbf{1}_{[s_h, s_h + c_h]}(\hat{s} - t) d\hat{s} \\ &= \int_{s_h - t}^{s_h + c_h - t} \mathbf{1}_{[s_h, s_h + c_h]}(\tau) d\tau \\ &= \begin{cases} c_h - t & t < c_h \\ 0 & t > c_h \end{cases} \end{aligned}$$

In particular, we have that for a.e. $h \in \mathbb{R}$ and for every $t \in (0,T)$ it holds

$$\partial_t v_h(t,s_h) = -\mathbf{1}_{[0,c_h]}(t).$$

Hence, for this particular solution, the 1D equation on the level set E_h is explicit:

$$\partial_t v_h + \partial_s v_h = c_h \mathbf{1}_{[0,c_h]},$$

which can be written also in the divergence form

$$\mathrm{div}_{t,s}(v_h(1,1)) = c_h \mathbf{1}_{[0,c_h]}. \tag{34}$$

From (34), we deduce immediately that, for a.e. $h \in \mathbb{R}$, the vector field $v_h(1,1)$ is a bounded, divergence-measure vector field in $(0,T)\times\mathbb{R}_s$. Applying Point 2 of Proposition 2 we can write for a.e. $t \in (0,T)$

$$v_h^-(t) - v_h^+(t) = +c_h \mathbf{1}_{[0,c_h]}(t) \tag{35}$$

where $v_h^\pm$ are the (L^∞ functions representing) Anzellotti traces on the surface $\Sigma_h := \{s = s_h\}$, defined as

$$v_h^\pm := \frac{\mathsf{Tr}^\pm\big(v_h(1,1),\Sigma_h\big)}{\mathsf{Tr}^\pm\big((1,1),\Sigma_h\big)} = \mathsf{Tr}^\pm\big(v_h(1,1),\Sigma_h\big).$$

We observe that by construction $v_h^- = 0$ a.e., hence (35) reduces to

$$-v_h^+ = c_h \mathbf{1}_{[0,c_h]}.$$

Taking now $\beta(\tau) = \tau^2$ and applying the Chain rule for Anzellotti traces (4) (being the vector field $v = (1,1)$ clearly of bounded variation) we obtain that for a.e. $h \in \mathbb{R}$ the vector field $w_h(1,1) := v_h^2(1,1)$ is still a divergence-measure vector field and it holds

$$w_h^- = 0, \qquad w_h^+ = +c_h^2 \mathbf{1}_{[0,c_h]},$$

i.e.

$$w_h^- - w_h^+ = -c_h^2 \mathbf{1}_{[0,c_h]}$$

so that, applying again Point 2 of Proposition 2, we can write

$$\mathrm{div}_{t,s}(w_h(1,1))\llcorner_{\Sigma_h} = (\partial_t \hat{w} + \partial_s \hat{w})\llcorner_{\Sigma_h} = -c_h^2 \mathbf{1}_{[0,c_h]}.$$

which in turn can be written as (recall $\hat{m}_h = c_h v_h \mathbf{1}_{s_h}$)

$$\partial_t \hat{w}_h + \partial_s \hat{w}_h = -c_h \partial_t (\hat{m}_h \mathscr{L}^1 \times \delta_{s_h}), \qquad \text{for a.e. } h \in \mathbb{R}.$$

Integrating and using Remark (1), we obtain the equation satisfied by u^2:

$$\partial_t(u^2) + \mathrm{div}(u^2 \mathbf{b}) = T,$$

being T the distribution defined by

$$T := -\partial_t(u^2 \mathbf{1}_S). \tag{36}$$

We conclude with the following

Proposition 4. *The distribution T defined in* (36) *is not representable by a Radon measure.*

Proof. By contradiction, assume that T is induced by some measure ξ: being the divergence of the bounded, measure-divergence vector field $w(1,\mathbf{b})$, we would necessarily have $\xi \ll \mathscr{H}^d$. On the other hand, it is immediate to see, directly from the construction of the Hamiltonian, that for any $\phi \in C_c^\infty, \|\phi\|_\infty \leq 1$ we have

$$\int_{[0,T]\times\mathbb{R}^2} \phi d\xi = \int_0^T \int_S u^2 \phi_t dt dx = \int_{\mathbb{R}} \int_0^T u(t,x_h) c_h^2 \phi_t(t,x_h)\, dt\, dh$$

which diverges being $c_h \notin L^2(\mathbb{R})$. Thus

$$\sup\left\{\langle T,\phi\rangle : \phi \in C_c^\infty((0,T)\times\mathbb{R}^2), \|\phi\|_\infty \leq 1\right\} = +\infty$$

which shows that T cannot be a distribution of order 0, hence it is not representable by a measure.

Acknowledgements The second author would like to thank Nikolay A. Gusev for bringing to his attention the reference [CGSW17] and for useful discussions on the topic.

References

[ABC13] G. Alberti, S. Bianchini, and G. Crippa. Structure of level sets and Sard-type properties of Lipschitz maps. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 12(4):863–902, 2013.

[ABC14] G. Alberti, S. Bianchini, and G. Crippa. A uniqueness result for the continuity equation in two dimensions. *J. Eur. Math. Soc. (JEMS)*, 16(2):201–234, 2014.

[ACM05] L. Ambrosio, G. Crippa, and S. Maniglia. Traces and fine properties of a *BD* class of vector fields and applications. *Ann. Fac. Sci. Toulouse Math. (6)*, 14(4):527–561, 2005.

[ADLM07] L. Ambrosio, C. De Lellis, and J. Malý. On the chain rule for the divergence of BV-like vector fields: applications, partial results, open problems. In *Perspectives in nonlinear partial differential equations*, volume 446 of *Contemp. Math.*, pages 31–67. Amer. Math. Soc., Providence, RI, 2007.

[ADM90] L. Ambrosio and G. Dal Maso. A general chain rule for distributional derivatives. *Proc. Amer. Math. Soc.*, 108(3):691–702, 1990.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Science Publications. Clarendon Press, 2000.

[Amb04] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Inventiones mathematicae*, 158(2):227–260, 2004.

[Anz83] G. Anzellotti. Traces of bounded vectorfields and the divergence theorem. 1983.

[BB17] S. Bianchini and P. Bonicatto. A uniqueness result for the decomposition of vector fields in $\mathbb{R}^d$. *Preprint*, 2017.

[BBG16] S. Bianchini, P. Bonicatto, and N. A. Gusev. Renormalization for autonomous nearly incompressible BV vector fields in two dimensions. *SIAM J. Math. Anal.*, 48(1):1–33, 2016.

[BG16] S. Bianchini and N. A. Gusev. Steady nearly incompressible vector fields in two-dimension: chain rule and renormalization. *Arch. Ration. Mech. Anal.*, 222(2):451–505, 2016.

[CGSW17] G. Crippa, N.A. Gusev, S. Spirito, and E. Wiedemann. Failure of the chain rule for the divergence of bounded vector fields. *Annali della Scuola Normale Superiore di Pisa*, XVII:1–18, 2017.

[DL89] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.

[DL07] C. De Lellis. Notes on hyperbolic systems of conservation laws and transport equations. In *Handbook of differential equations: evolutionary equations. Vol. III*, Handb. Differ. Equ., pages 277–382. Elsevier/North-Holland, Amsterdam, 2007.

[Kru70] S. N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.

[Vol67] A. I. Volpert. Spaces BV and quasilinear equations. *Mat. Sb. (N.S.)*, 73 (115):255–302, 1967.