Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields

Luigi Ambrosio, Gianluca Crippa

Scuola Normale Superiore
Piazza dei Cavalieri 7, 56126 Pisa, Italy
l.ambrosio@sns.it, g.crippa@sns.it

Contents

1 Introduction

In these notes we discuss some recent progress on the problem of the existence, uniqueness and stability of the flow associated to a weakly differentiable (Sobolev or BV regularity with respect to the spatial variables) time-dependent vector field $\boldsymbol{b}(t,x) = \boldsymbol{b}_t(x)$ in $\mathbb{R}^d$. Vector fields with this "low" regularity show up, for instance, in several PDE's describing the motion of fluids, and in the theory of conservation laws.

We are therefore interested to the well posedness of the system of ordinary differential equations

$$\text{(ODE)} \qquad \begin{cases} \dot\gamma(t) = \boldsymbol{b}_t(\gamma(t)) \\ \gamma(0) \ = \ x. \end{cases}$$

In some situations one might hope for a "generic" uniqueness of the solutions of (ODE), i.e. for "almost every" initial datum x. But, as a matter of fact, no such uniqueness theorem is presently known in the case when $\boldsymbol{b}(t,\cdot)$ has a Sobolev or BV regularity (this issue is discussed in §9).

An even weaker requirement is the research of a "canonical selection principle", i.e. a strategy to select for $\mathscr{L}^d$-almost every x a solution $\boldsymbol{X}(\cdot,x)$ in such a way that this selection is stable w.r.t. smooth approximations of $\boldsymbol{b}$.

In other words, we would like to know that, whenever we approximate $\boldsymbol{b}$ by smooth vector fields $\boldsymbol{b}^h$, the classical trajectories $\boldsymbol{X}^h$ associated to $\boldsymbol{b}^h$ satisfy

$$\lim_{h\to\infty} \boldsymbol{X}^h(\cdot,x) = \boldsymbol{X}(\cdot,x) \qquad \text{in } C([0,T];\mathbb{R}^d), \text{ for } \mathscr{L}^d\text{-a.e. } x.$$

The following simple example, borrowed from [8], provides an illustration of the kind of phenomena that can occur.

Example 1 Let us consider the autonomous ODE

$$\begin{cases} \dot\gamma(t) = \sqrt{|\gamma(t)|} \\ \gamma(0) \ = \ x_0. \end{cases}$$

Then, solutions of the ODE are not unique for $x_0 = -c^2 < 0$. Indeed, they reach the origin in time $2c$, where they can stay for an arbitrary time T, then continuing as $x(t) = \frac{1}{4}(t-T-2c)^2$.

Let us consider for instance the Lipschitz approximation (that could easily be made smooth) of $b(\gamma) = \sqrt{|\gamma|}$ by

$$b_\epsilon(\gamma) := \begin{cases} \sqrt{|\gamma|} & \text{if } -\infty < \gamma \le -\epsilon^2; \\ \epsilon & \text{if } -\epsilon^2 \le \gamma \le \lambda_\epsilon - \epsilon^2 \\ \sqrt{\gamma - \lambda_\epsilon + 2\epsilon^2} & \text{if } \lambda_\epsilon - \epsilon^2 \le \gamma < +\infty, \end{cases}$$

with $\lambda_\epsilon - \epsilon^2 > 0$. Then, solutions of the approximating ODE's starting from $-c^2$ reach the value $-\epsilon^2$ in time $t_\epsilon = 2(c-\epsilon)$ and then they continue with constant speed ϵ until they reach $\lambda_\epsilon - \epsilon^2$, in time $T_\epsilon = \lambda_\epsilon/\epsilon$. Then, they continue as $\lambda_\epsilon - 2\epsilon^2 + \frac{1}{4}(t - t_\epsilon - T_\epsilon)^2$.

Choosing $\lambda_\epsilon = \epsilon T$, with $T > 0$, by this approximation we select the solutions that don't move, when at the origin, exactly for a time T.

Other approximations, as for instance $b_\epsilon(\gamma) = \sqrt{\epsilon + |\gamma|}$, select the solutions that move immediately away from the singularity at $\gamma = 0$. Among all possibilities, this family of solutions

$x(t, x_0)$ is singled out by the property that $x(t,\cdot)_\# \mathscr{L}^1$ is absolutely continuous with respect to $\mathscr{L}^1$, so no concentration of trajectories occurs at the origin. To see this fact, notice that we can integrate in time the identity

$$0 = x(t,\cdot)_\# \mathscr{L}^1(\{0\}) = \mathscr{L}^1\left(\{x_0 : \ x(t,x_0) = 0\}\right)$$

and use Fubini's theorem to obtain

$$0 = \int \mathscr{L}^1(\{t : \ x(t,x_0) = 0\})\, dx_0.$$

Hence, for $\mathscr{L}^1$-a.e. x_0, $x(\cdot, x_0)$ does not stay at 0 for a strictly positive set of times.

The theme of existence, uniqueness and stability has been trated in detail in the lectures notes [7], and more recently in [8] (where also the applications to systems of conservation laws [11], [9] and to the semi-geostrophic equation [53] are described), so some overlap with the content of these notes is unavoidable. Because of this fact, we decided to put here more emphasis on even more recent results [73], [15], [17], [48], relative to the differentiability properties of $\boldsymbol{X}(t,x)$ with respect to x. This is not a casual choice, as the key idea of the paper [15] was found during the Bologna school.

2 The continuity equation

An important tool, in studying existence, uniqueness ad stability of (ODE), is the well-posedness of the Cauchy problem for the homogeneous conservative continuity equation

$$\text{(PDE)} \qquad \frac{d}{dt}\mu_t + D_x \cdot (\boldsymbol{b}\mu_t) = 0 \qquad (t,x) \in I \times \mathbb{R}^d$$

and for the transport equation

$$\frac{d}{dt} w_t + \boldsymbol{b} \cdot \nabla w_t = c_t.$$

We will see that there is a close link between (PDE) and (ODE), first investigated in a nonsmooth setting by DiPerna and Lions in [61].
Let us now make some basic technical remarks on the continuity equation and the transport equation:

Remark 2 (Regularity in space of $\boldsymbol{b}_t$ and μ_t) (1) Since the continuity equation (PDE) is in divergence form, it makes sense without *any* regularity requirement on $\boldsymbol{b}_t$ and/or μ_t, provided

$$\int_I \int_A |\boldsymbol{b}_t|\, d|\mu_t|\, dt < +\infty \qquad \forall A \subset\subset \mathbb{R}^d. \tag{1}$$

However, when we consider possibly singular measures μ_t, we must take care of the fact that the product $\boldsymbol{b}_t \mu_t$ is sensitive to modifications of $\boldsymbol{b}_t$ in $\mathscr{L}^d$-negligible sets. In the Sobolev or BV case we will consider only measures $\mu_t = w_t \mathscr{L}^d$, so everything is well stated.

(2) On the other hand, due to the fact that the distribution $\boldsymbol{b}_t \cdot \nabla w$ is defined by

$$\langle \boldsymbol{b}_t \cdot \nabla w, \varphi \rangle := -\int_I \int w \boldsymbol{b}_t \cdot \nabla \varphi \, dx dt - \int_I \langle D_x \cdot \boldsymbol{b}_t, w_t \varphi_t \rangle \, dt \qquad \forall \varphi \in C_c^\infty(I \times \mathbb{R}^d)$$

(a definition consistent with the case when w_t is smooth) the transport equation makes sense *only* if we assume that $D_x \cdot \boldsymbol{b}_t = \operatorname{div} \boldsymbol{b}_t \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in I$. See also [28], [29] for recent results on the transport equation when $\boldsymbol{b}$ satisfies a one-sided Lipschitz condition.

Next, we consider the problem of the time continuity of $t \mapsto \mu_t$ and $t \mapsto w_t$.

Remark 3 (Regularity in time of μ_t) For any test function $\varphi \in C_c^\infty(\mathbb{R}^d)$, condition (1) gives

$$\frac{d}{dt} \int_{\mathbb{R}^d} \varphi \, d\mu_t = \int_{\mathbb{R}^d} \boldsymbol{b}_t \cdot \nabla \varphi \, d\mu_t \in L^1(I)$$

and therefore the map $t \mapsto \langle \mu_t, \varphi \rangle$, for given φ, has a unique uniformly continuous representative in I. By a simple density argument we can find a unique representative $\tilde{\mu}_t$ independent of φ, such that $t \mapsto \langle \tilde{\mu}_t, \varphi \rangle$ is uniformly continuous in I for any $\varphi \in C_c^\infty(\mathbb{R}^d)$. We will always work with this representative, so that μ_t will be well defined *for all* t and even at the endpoints of I. An analogous remark applies for solutions of the transport equation.

There are some other important links between the two equations:
(1) The transport equation reduces to the continuity equation in the case when $c_t = -w_t \operatorname{div} \boldsymbol{b}_t$.
(2) Formally, one can estabilish a duality between the two equations via the (formal) identity

$$\begin{aligned} \frac{d}{dt} \int w_t \, d\mu_t &= \int \frac{d}{dt} w_t \, d\mu_t + \int \frac{d}{dt} \mu_t w_t \\ &= \int (-\boldsymbol{b}_t \cdot \nabla w_t + c_t) \, d\mu_t + \int \boldsymbol{b}_t \cdot \nabla w_t \, d\mu_t = \int c_t \, d\mu_t. \end{aligned}$$

This duality method is a classical tool to prove uniqueness in a sufficiently smooth setting (but see also [28], [29]).
(3) Finally, if we denote by $\boldsymbol{Y}(t,s,x)$ the solution of the ODE at time t, starting from x at the initial time s, *i.e.*

$$\frac{d}{dt} \boldsymbol{Y}(t,s,x) = \boldsymbol{b}_t(\boldsymbol{Y}(t,s,x)), \qquad \boldsymbol{Y}(s,s,x) = x,$$

then $\boldsymbol{Y}(t,\cdot,\cdot)$ are themselves solutions of the transport equation: to see this, it suffices to differentiate the semigroup identity

$$\boldsymbol{Y}(t,s,\boldsymbol{Y}(s,l,x)) = \boldsymbol{Y}(t,l,x)$$

w.r.t. s to obtain, after the change of variables $y = \boldsymbol{Y}(s,l,x)$, the equation

$$\frac{d}{ds} \boldsymbol{Y}(t,s,y) + \boldsymbol{b}_s(y) \cdot \nabla \boldsymbol{Y}(t,s,y) = 0.$$

This property is used in a essential way in [61] to characterize the flow $\boldsymbol{Y}$ and to prove its stability properties. The approach developed here, based on [6], is based on a careful analysis of the measures transported by the flow, and ultimately on the homogeneous continuity equation only.

3 The continuity equation within the Cauchy-Lipschitz framework

In this section we recall the classical representation formulas for solutions of the continuity or transport equation in the case when

$$\boldsymbol{b} \in L^1\left([0,T]; W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)\right).$$

Under this assumption it is well known that solutions $\boldsymbol{X}(t,\cdot)$ of the ODE are unique and stable. A quantitative information can be obtained by differentiation:

$$\begin{aligned}\frac{d}{dt}|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|^2 &= 2\langle \boldsymbol{b}_t(\boldsymbol{X}(t,x))-\boldsymbol{b}_t(\boldsymbol{X}(t,y)), \boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)\rangle \\ &\leq 2\mathrm{Lip}\,(\boldsymbol{b}_t)|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|^2\end{aligned}$$

(here $\mathrm{Lip}\,(f)$ denotes the least Lipschitz constant of f), so that Gronwall lemma immediately gives

$$\mathrm{Lip}\,(\boldsymbol{X}(t,\cdot)) \leq \exp\left(\int_0^t \mathrm{Lip}\,(\boldsymbol{b}_s)\,ds\right). \tag{2}$$

Turning to the continuity equation, uniqueness of measure-valued solutions can be proved by the duality method. Or, following the techniques developed in these lectures, it can be proved in a more general setting for positive measure-valued solutions (via the superposition principle) and for signed solutions $\mu_t = w_t\mathscr{L}^d$ (via the theory of renormalized solutions). So in this section we focus only on the existence and the representation issues.
The representation formula is indeed very simple:

Proposition 4 *For any initial datum $\bar\mu$ the solution of the continuity equation is given by*

$$\mu_t := \boldsymbol{X}(t,\cdot)_\#\bar\mu, \quad i.e. \quad \int_{\mathbb{R}^d}\varphi\,d\mu_t = \int_{\mathbb{R}^d}\varphi(\boldsymbol{X}(t,x))\,d\bar\mu(x) \quad \forall \varphi\in C_b(\mathbb{R}^d). \tag{3}$$

Proof. Notice first that we need only to check the distributional identity $\frac{d}{dt}\mu_t + D_x\cdot(\boldsymbol{b}_t\mu_t)=0$ on test functions of the form $\psi(t)\varphi(x)$, so that

$$\int_{\mathbb{R}}\psi'(t)\langle\mu_t,\varphi\rangle\,dt + \int_{\mathbb{R}}\psi(t)\int_{\mathbb{R}^d}\langle\boldsymbol{b}_t,\nabla\varphi\rangle\,d\mu_t\,dt = 0.$$

This means that we have to check that $t\mapsto\langle\mu_t,\varphi\rangle$ belongs to $W^{1,1}(0,T)$ for any $\varphi\in C_c^\infty(\mathbb{R}^d)$ and that its distributional derivative is $\int_{\mathbb{R}^d}\langle\boldsymbol{b}_t,\nabla\varphi\rangle\,d\mu_t$.
We show first that this map is absolutely continuous, and in particular $W^{1,1}(0,T)$; then one needs only to compute the pointwise derivative. For every choice of finitely many, say n, pairwise disjoint intervals $(a_i,b_i)\subset[0,T]$ we have

$$\begin{aligned}\sum_{i=1}^n|\varphi(\boldsymbol{X}(b_i,x))-\varphi(\boldsymbol{X}(a_i,x))| &\leq \|\nabla\varphi\|_\infty\int_{\cup_i(a_i,b_i)}|\dot{\boldsymbol{X}}(t,x)|\,dt \\ &\leq \|\nabla\varphi\|_\infty\int_{\cup_i(a_i,b_i)}\sup|\boldsymbol{b}_t|\,dt\end{aligned}$$

and therefore an integration with respect to $\bar\mu$ gives

$$\sum_{i=1}^n |\langle \mu_{b_i} - \mu_{a_i}, \varphi\rangle| \le \bar\mu(\mathbb{R}^d)\|\nabla\varphi\|_\infty \int_{\cup_i(a_i,b_i)} \sup|\boldsymbol{b}_t|\,dt.$$

The absolute continuity of the integral shows that the right hand side can be made small when $\sum_i(b_i - a_i)$ is small. This proves the absolute continuity. For any x the identity $\dot{\boldsymbol{X}}(t,x) = \boldsymbol{b}_t(\boldsymbol{X}(t,x))$ is fulfilled for $\mathscr{L}^1$-a.e. $t \in [0,T]$. Then, by Fubini's theorem, we know also that for $\mathscr{L}^1$-a.e. $t \in [0,T]$ the previous identity holds for $\bar\mu$-a.e. x, and therefore

$$\begin{aligned}
\frac{d}{dt}\langle \mu_t, \varphi\rangle &= \frac{d}{dt}\int_{\mathbb{R}^d} \varphi(\boldsymbol{X}(t,x))\,d\bar\mu(x) \\
&= \int_{\mathbb{R}^d} \langle \nabla\varphi(\boldsymbol{X}(t,x)), \boldsymbol{b}_t(\boldsymbol{X}(t,x))\rangle\,d\bar\mu(x) \\
&= \langle \boldsymbol{b}_t\mu_t, \nabla\varphi\rangle
\end{aligned}$$

for $\mathscr{L}^1$-a.e. $t \in [0,T]$. □

In the case when $\bar\mu = \rho\mathscr{L}^d$ we can say something more, proving that the measures $\mu_t = \boldsymbol{X}(t,\cdot)_\#\bar\mu$ are absolutely continuous w.r.t. $\mathscr{L}^d$ and computing *explicitely* their density. Let us start by recalling the classical *area formula*: if $f : \mathbb{R}^d \to \mathbb{R}^d$ is a (locally) Lipschitz map, then

$$\int_A g|Jf|\,dx = \int_{\mathbb{R}^d} \sum_{x\in A\cap f^{-1}(y)} g(x)\,dy$$

for any Borel set $A \subset \mathbb{R}^d$ and any integrable function $g : A \to \mathbb{R}$, where $Jf = \det\nabla f$ (recall that, by Rademacher theorem, Lipschitz functions are differentiable $\mathscr{L}^d$-a.e.). Assuming in addition that f is one to one and onto and that $|Jf| > 0$ $\mathscr{L}^d$-a.e. on A we can set $A = f^{-1}(B)$ and $g = \rho/|Jf|$ to obtain

$$\int_{f^{-1}(B)} \rho\,dx = \int_B \frac{\rho}{|Jf|}\circ f^{-1}\,dy.$$

In other words, we have got a formula for the push-forward:

$$f_\#(\rho\mathscr{L}^d) = \frac{\rho}{|Jf|}\circ f^{-1}\mathscr{L}^d. \tag{4}$$

In our case $f(x) = \boldsymbol{X}(t,x)$ is surely one to one, onto and Lipschitz. It remains to show that $|J\boldsymbol{X}(t,\cdot)|$ does not vanish: in fact, one can show that $J\boldsymbol{X} > 0$ and

$$\exp\left[-\int_0^t \|[\mathrm{div}\,\boldsymbol{b}_s]^-\|_\infty\,ds\right] \le J\boldsymbol{X}(t,x) \le \exp\left[\int_0^t \|[\mathrm{div}\,\boldsymbol{b}_s]^+\|_\infty\,ds\right] \tag{5}$$

for $\mathscr{L}^d$-a.e. x, thanks to the following fact, whose proof is left as an exercise.

Exercise 5 If $\boldsymbol{b}$ is smooth, we have

$$\frac{d}{dt} J\boldsymbol{X}(t,x) = \operatorname{div} \boldsymbol{b}_t(\boldsymbol{X}(t,x)) J\boldsymbol{X}(t,x).$$

Hint: use the ODE $\frac{d}{dt}\nabla \boldsymbol{X} = \nabla \boldsymbol{b}_t(\boldsymbol{X})\nabla \boldsymbol{X}$.

The previous exercise gives that, in the smooth case, $J\boldsymbol{X}(\cdot,x)$ solves a linear ODE with the initial condition $J\boldsymbol{X}(0,x) = 1$, whence the estimates on $J\boldsymbol{X}$ follow. In the general case the upper estimate on $J\boldsymbol{X}$ still holds by a smoothing argument, thanks to the lower semicontinuity of

$$\Phi(v) := \begin{cases} \|Jv\|_\infty & \text{if } Jv \geq 0 \ \mathscr{L}^d\text{-a.e.} \\ \\ +\infty & \text{otherwise} \end{cases}$$

with respect to the w^*-topology of $W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$. This is indeed the supremum of the family of $\Phi_p^{1/p}$, where Φ_p are the *polyconvex* (and therefore lower semicontinuous) functionals

$$\Phi_p(v) := \int_{B_p} |\chi(Jv)|^p \, dx.$$

Here $\chi(t)$, equal to ∞ on $(-\infty,0)$ and equal to t on $[0,+\infty)$, is l.s.c. and convex. The lower estimate can be obtained by applying the upper one in a time reversed situation.
Now we turn to the representation of solutions of the transport equation:

Proposition 6 *If $w \in L^1_{\mathrm{loc}}\left([0,T]\times\mathbb{R}^d\right)$ solves*

$$\frac{d}{dt} w_t + \boldsymbol{b}\cdot\nabla w_t = c \in L^1_{\mathrm{loc}}\left([0,T]\times\mathbb{R}^d\right)$$

then, for $\mathscr{L}^d$-a.e. x, we have

$$w_t(\boldsymbol{X}(t,x)) = w_0(x) + \int_0^t c_s(\boldsymbol{X}(s,x))\, ds \qquad \forall t \in [0,T].$$

The (formal) proof is based on the simple observation that

$$\begin{aligned} \frac{d}{dt} w_t \circ \boldsymbol{X}(t,x) &= \frac{d}{dt} w_t(\boldsymbol{X}(t,x)) + \frac{d}{dt}\boldsymbol{X}(t,x)\cdot\nabla w_t(\boldsymbol{X}(t,x)) \\ &= \frac{d}{dt} w_t(\boldsymbol{X}(t,x)) + \boldsymbol{b}_t(\boldsymbol{X}(t,x))\cdot\nabla w_t(\boldsymbol{X}(t,x)) \\ &= c_t(\boldsymbol{X}(t,x)). \end{aligned}$$

In particular, as $\boldsymbol{X}(t,x) = \boldsymbol{Y}(t,0,x) = [\boldsymbol{Y}(0,t,\cdot)]^{-1}(x)$, we get

$$w_t(y) = w_0(\boldsymbol{Y}(0,t,y)) + \int_0^t c_s(\boldsymbol{Y}(s,t,y))\, ds.$$

We conclude this presentation of the classical theory pointing out two simple local variants of the assumption $\boldsymbol{b} \in L^1\left([0,T]; W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)\right)$ made throughout this section.

Remark 7 (First local variant) The theory outlined above still works under the assumptions

$$\boldsymbol{b} \in L^1\left([0,T]; W^{1,\infty}_{\text{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right), \qquad \frac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right).$$

Indeed, due to the growth condition on $\boldsymbol{b}$, we still have pointwise uniqueness of the ODE and a uniform local control on the growth of $|\boldsymbol{X}(t,x)|$, therefore we need only to consider a *local* Lipschitz condition w.r.t. x, integrable w.r.t. t.

The next variant will be used in the proof of the superposition principle.

Remark 8 (Second local variant) Still keeping the $L^1(W^{1,\infty}_{\text{loc}})$ assumption, and assuming $\mu_t \geq 0$, the second growth condition on $|\boldsymbol{b}|$ can be replaced by a global, but more intrinsic, condition:

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_t|}{1+|x|}\, d\mu_t\, dt < +\infty. \tag{6}$$

Under this assumption one can show that for $\bar{\mu}$-a.e. x the *maximal* solution $\boldsymbol{X}(\cdot,x)$ of the ODE starting from x is defined up to $t = T$ and still the representation $\mu_t = \boldsymbol{X}(t,\cdot)_{\#}\bar{\mu}$ holds for $t \in [0,T]$.

4 (ODE) uniqueness versus (PDE) uniqueness

In this section we illustrate some quite general principles, whose application may depend on specific assumptions on $\boldsymbol{b}$, relating the uniqueness for the ODE to the uniqueness for the PDE. The viewpoint adopted in this section is very close in spirit to Young's theory [95] of generalized surfaces and controls (a theory with remarkable applications also to non-linear PDE's [60, 88] and to Calculus of Variations [19]) and has also some connection with Brenier's weak solutions of incompressible Euler equations [30], with Kantorovich's viewpoint in the theory of optimal transportation [63, 85] and with Mather's theory [80, 81, 20]: in order to study existence, uniqueness and stability with respect to perturbations of the data of solutions to the ODE, we consider suitable measures in the space of continuous maps, allowing for superposition of trajectories. Then, in some special situations we are able to show that this superposition actually does not occur, but still this "probabilistic" interpretation is very useful to understand the underlying techniques and to give an intrinsic characterization of the flow.
The first very general criterion is the following.

Theorem 9 *Let $A \subset \mathbb{R}^d$ be a Borel set. The following two properties are equivalent:*

(a) Solutions of the ODE are unique for any $x \in A$.

(b) Nonnegative measure-valued solutions of the PDE are unique for any $\bar{\mu}$ concentrated in A, i.e. such that $\bar{\mu}(\mathbb{R}^d \setminus A) = 0$.

Proof. It is clear that (b) implies (a), just choosing $\bar{\mu} = \delta_x$ and noticing that two different solutions $\boldsymbol{X}(t)$, $\tilde{\boldsymbol{X}}(t)$ of the ODE induce two different solutions of the PDE, namely $\delta_{\boldsymbol{X}(t)}$ and $\delta_{\tilde{\boldsymbol{X}}(t)}$.
The converse implication is less obvious and requires the superposition principle that we are going to describe below, and that provides the representation

$$\int_{\mathbb{R}^d} \varphi \, d\mu_t = \int_{\mathbb{R}^d} \left(\int_{\Gamma_T} \varphi(\gamma(t)) \, d\boldsymbol{\eta}_x(\gamma) \right) d\mu_0(x) \qquad \forall \varphi \in C_b(\mathbb{R}^d),$$

with $\boldsymbol{\eta}_x$ probability measures concentrated on the absolutely continuous integral solutions of the ODE starting from x. Therefore, when these are unique, the measures $\boldsymbol{\eta}_x$ are unique (and are Dirac masses), so that the solutions of the PDE are unique. □

We will use the shorter notation Γ_T for the space $C\left([0,T];\mathbb{R}^d\right)$ and denote by $e_t : \Gamma_T \to \mathbb{R}^d$ the evaluation maps $\gamma \mapsto \gamma(t)$, $t \in [0,T]$.

Definition 10 (Superposition solutions) *Let $\boldsymbol{\eta} \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ be a measure concentrated on the set of pairs (x,γ) such that γ is an absolutely continuous integral solution of the ODE with $\gamma(0) = x$. We define*

$$\langle \mu_t^{\boldsymbol{\eta}}, \varphi \rangle := \int_{\mathbb{R}^d \times \Gamma_T} \varphi(e_t(\gamma)) \, d\boldsymbol{\eta}(x,\gamma) \qquad \forall \varphi \in C_b(\mathbb{R}^d).$$

By a standard approximation argument the identity defining $\mu_t^{\boldsymbol{\eta}}$ holds for any Borel function φ such that $\gamma \mapsto \varphi(e_t(\gamma))$ is $\boldsymbol{\eta}$-integrable (or equivalently for any $\mu_t^{\boldsymbol{\eta}}$-integrable function φ).
Under the (local) integrability condition

$$\int_0^T \int_{\mathbb{R}^d \times \Gamma_T} \chi_{B_R}(e_t) |\boldsymbol{b}_t(e_t)| \, d\boldsymbol{\eta} \, dt < +\infty \qquad \forall R > 0 \tag{7}$$

it is not hard to see that $\mu_t^{\boldsymbol{\eta}}$ solves the PDE with the initial condition $\bar{\mu} := (\pi_{\mathbb{R}^d})_\# \boldsymbol{\eta}$: indeed, let us check first that $t \mapsto \langle \mu_t^{\boldsymbol{\eta}}, \varphi \rangle$ is absolutely continuous for any $\varphi \in C_c^\infty(\mathbb{R}^d)$. For every choice of finitely many pairwise disjoint intervals $(a_i, b_i) \subset [0,T]$ we have

$$\sum_{i=1}^n |\varphi(\gamma(b_i)) - \varphi(\gamma(a_i))| \le \mathrm{Lip}\,(\varphi) \int_{\cup_i (a_i,b_i)} \chi_{B_R}(|e_t(\gamma)|) \boldsymbol{b}_t(e_t(\gamma))| \, dt$$

for $\boldsymbol{\eta}$-a.e. (x,γ), with R such that $\operatorname{supp} \varphi \subset \overline{B}_R$. Therefore an integration with respect to $\boldsymbol{\eta}$ gives

$$\sum_{i=1}^n |\langle \mu_{b_i}^{\boldsymbol{\eta}}, \varphi \rangle - \langle \mu_{a_i}^{\boldsymbol{\eta}}, \varphi \rangle| \le \mathrm{Lip}\,(\varphi) \int_{\cup_i (a_i,b_i)} \int_{\mathbb{R}^d \times \Gamma_T} \chi_{B_R}(e_t) |\boldsymbol{b}_t(e_t)| \, d\boldsymbol{\eta} \, dt.$$

The absolute continuity of the integral shows that the right hand side can be made small when $\sum_i (b_i - a_i)$ is small. This proves the absolute continuity.

It remains to evaluate the time derivative of $t \mapsto \langle \mu_t^{\boldsymbol{\eta}}, \varphi \rangle$: we know that for $\boldsymbol{\eta}$-a.e. (x, γ) the identity $\dot\gamma(t) = \boldsymbol{b}_t(\gamma(t))$ is fulfilled for $\mathscr{L}^1$-a.e. $t \in [0, T]$. Then, by Fubini's theorem, we know also that for $\mathscr{L}^1$-a.e. $t \in [0, T]$ the previous identity holds for $\boldsymbol{\eta}$-a.e. (x, γ), and therefore

$$\begin{aligned}\frac{d}{dt}\langle \mu_t^{\boldsymbol{\eta}}, \varphi\rangle &= \frac{d}{dt}\int_{\mathbb{R}^d\times\Gamma_T} \varphi(e_t(\gamma))\, d\boldsymbol{\eta} \\ &= \int_{\mathbb{R}^d\times\Gamma_T} \langle \nabla\varphi(e_t(\gamma)), \boldsymbol{b}_t(e_t(\gamma))\rangle\, d\boldsymbol{\eta} = \langle \boldsymbol{b}_t\mu_t, \nabla\varphi\rangle \quad \mathscr{L}^1\text{-a.e. in } [0,T].\end{aligned}$$

Remark 11 Actually the formula defining $\mu_t^{\boldsymbol{\eta}}$ does not contain x, and so it involves only the projection of $\boldsymbol{\eta}$ on Γ_T. Therefore one could also consider measures $\boldsymbol{\sigma}$ in Γ_T, concentrated on the set of solutions of the ODE (for an arbitrary initial point x). These two viewpoints are basically equivalent: given $\boldsymbol{\eta}$ one can build $\boldsymbol{\sigma}$ just by projection on Γ_T, and given $\boldsymbol{\sigma}$ one can consider the conditional probability measures $\boldsymbol{\eta}_x$ concentrated on the solutions of the ODE starting from x induced by the random variable $\gamma \mapsto \gamma(0)$ in Γ_T, the law $\bar\mu$ (i.e. the push forward) of the same random variable and recover $\boldsymbol{\eta}$ as follows:

$$\int_{\mathbb{R}^d\times\Gamma_T} \varphi(x,\gamma)\, d\boldsymbol{\eta}(x,\gamma) := \int_{\mathbb{R}^d}\left(\int_{\Gamma_T}\varphi(x,\gamma)\, d\boldsymbol{\eta}_x(\gamma)\right) d\bar\mu(x). \tag{8}$$

Our viewpoint has been chosen just for technical convenience, to avoid the use, wherever this is possible, of the conditional probability theorem.

By restricting $\boldsymbol{\eta}$ to suitable subsets of $\mathbb{R}^d\times\Gamma_T$, several manipulations with superposition solutions of the continuity equation are possible and useful, and these are not immediate to see just at the level of general solutions of the continuity equation. This is why the following result is interesting.

Theorem 12 (Superposition principle) *Let $\mu_t \in \mathscr{M}_+(\mathbb{R}^d)$ be a solution of (PDE) and assume that*

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}|_t(x)}{1+|x|}\, d\mu_t\, dt < +\infty.$$

Then μ_t is a superposition solution, i.e. there exists $\boldsymbol{\eta} \in \mathscr{M}_+(\mathbb{R}^d\times\Gamma_T)$ such that $\mu_t = \mu_t^{\boldsymbol{\eta}}$ for any $t \in [0,T]$.

In the proof we use the *narrow* convergence of positive measures, i.e. the convergence with respect to the duality with continuous and bounded functions, and the *easy* implication in Prokhorov compactness theorem: any tight and bounded family $\mathscr{F}$ in $\mathscr{M}_+(X)$ is (sequentially) relatively compact w.r.t. the narrow convergence. Remember that tightness means:

$$\text{for any } \epsilon > 0 \text{ there exists } K \subset X \text{ compact s.t. } \mu(X\setminus K) < \epsilon\ \forall \mu \in \mathscr{F}.$$

A necessary and sufficient condition for tightness is the existence of a *coercive* functional $\Psi : X \to [0,\infty]$ such that $\int \Psi\, d\mu \le 1$ for any $\mu \in \mathscr{F}$.

Proof. **Step 1.** (Smoothing) The smoothing argument which follows has been inspired by [65]. We mollify μ_t w.r.t. the space variable with a kernel ρ having finite first moment M and support equal to the whole of $\mathbb{R}^d$ (a Gaussian, for instance), obtaining smooth and strictly positive functions μ_t^ϵ. We also choose a function $\psi : \mathbb{R}^d \to [0, +\infty)$ such that $\psi(x) \to +\infty$ as $|x| \to +\infty$ and

$$\int_{\mathbb{R}^d} \psi(x) \mu_0 * \rho_\epsilon(x)\, dx \le 1 \qquad \forall \epsilon \in (0,1)$$

and a convex nondecreasing function $\Theta : \mathbb{R}^+ \to \mathbb{R}$ having a more than linear growth at infinity such that

$$\int_0^T \int_{\mathbb{R}^d} \frac{\Theta(|\boldsymbol{b}_t|(x))}{1+|x|}\, d\mu_t dt < +\infty$$

(the existence of Θ is ensured by Dunford-Pettis theorem). Defining

$$\mu_t^\epsilon := \mu_t * \rho_\epsilon, \qquad \boldsymbol{b}_t^\epsilon := \frac{(\boldsymbol{b}_t \mu_t) * \rho_\epsilon}{\mu_t^\epsilon},$$

it is immediate that

$$\frac{d}{dt}\mu_t^\epsilon + D_x \cdot (\boldsymbol{b}_t^\epsilon \mu_t^\epsilon) = \frac{d}{dt}\mu_t * \rho_\epsilon + D_x \cdot (\boldsymbol{b}_t \mu_t) * \rho_\epsilon = 0$$

and that $\boldsymbol{b}^\epsilon \in L^1\left([0,T]; W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right)$. Therefore Remark 8 can be applied and the representation $\mu_t^\epsilon = \boldsymbol{X}^\epsilon(t,\cdot)_\# \mu_0^\epsilon$ still holds. Then, we define

$$\boldsymbol{\eta}^\epsilon := (x, \boldsymbol{X}^\epsilon(\cdot, x))_\# \, \mu_0^\epsilon,$$

so that

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi \, d\mu_t^{\boldsymbol{\eta}_\epsilon} &= \int_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t)) \, d\boldsymbol{\eta}^\epsilon \\ &= \int_{\mathbb{R}^d} \varphi(\boldsymbol{X}^\epsilon(t,x)) \, d\mu_0^\epsilon(x) = \int_{\mathbb{R}^d} \varphi \, d\mu_t^\epsilon. \end{aligned} \tag{9}$$

Step 2. (Tightness) We will be using the inequality

$$((1+|x|)c) * \rho_\epsilon \le (1+|x|) c * \rho_\epsilon + \epsilon c * \tilde{\rho}_\epsilon \tag{10}$$

for c nonnegative measure and $\tilde{\rho}(y) = |y|\rho(y)$, and

$$\Theta(|\boldsymbol{b}_t^\epsilon(x)|)\mu_t^\epsilon(x) \le (\Theta(|\boldsymbol{b}_t|)\mu_t) * \rho_\epsilon(x). \tag{11}$$

The proof of the first one is elementary, while the proof of the second one follows by applying Jensen's inequality with the convex l.s.c. function $(z,t) \mapsto \Theta(|z|/t)t$ (set equal to $+\infty$ if $t<0$, or $t=0$ and $z \neq 0$, and equal to 0 if $z=t=0$) and with the measure $\rho_\epsilon(x-\cdot)\mathscr{L}^d$.

Let us introduce the functional

$$\Psi(x,\gamma) := \psi(x) + \int_0^T \frac{\Theta(|\dot\gamma|)}{1+|\gamma|}\,dt,$$

set equal to $+\infty$ on $\Gamma_T \setminus AC([0,T];\mathbb{R}^d)$.
Using Ascoli-Arzelá theorem, it is not hard to show that Ψ is coercive (it suffices to show that $\max|\gamma|$ is bounded on the sublevels $\{\Psi \le t\}$). Since

$$\begin{aligned}\int_{\mathbb{R}^d\times\Gamma_T}\int_0^T \frac{\Theta(|\dot\gamma|)}{1+|\gamma|}\,dt\,d\boldsymbol{\eta}^\epsilon(x,\gamma) &= \int_0^T\int_{\mathbb{R}^d}\frac{\Theta(|\boldsymbol{b}_t^\epsilon|)}{1+|x|}\,d\mu_t^\epsilon\,dt\\ &\overset{(10),(11)}{\le} (1+\epsilon M)\int_0^T\int_{\mathbb{R}^d}\frac{\Theta(|\boldsymbol{b}_t|(x))}{1+|x|}\,d\mu_t dt\end{aligned}$$

and

$$\int_{\mathbb{R}^d\times\Gamma_T}\psi(x)\,d\boldsymbol{\eta}^\epsilon(x,\gamma) = \int_{\mathbb{R}^d}\psi(x)\,d\mu_0^\epsilon \le 1$$

we obtain that $\int \Psi\,d\boldsymbol{\eta}^\epsilon$ is uniformly bounded for $\epsilon\in(0,1)$, and therefore Prokhorov compactness theorem tells us that the family $\boldsymbol{\eta}^\epsilon$ is narrowly sequentially relatively compact as $\epsilon \downarrow 0$. If $\boldsymbol{\eta}$ is any limit point we can pass to the limit in (9) to obtain that $\mu_t = \mu_t^{\boldsymbol{\eta}}$.
Step 3. ($\boldsymbol{\eta}$ is concentrated on solutions of the ODE) It suffices to show that

$$\int_{\mathbb{R}^d\times\Gamma_T}\frac{\left|\gamma(t)-x-\int_0^t \boldsymbol{b}_s(\gamma(s))\,ds\right|}{1+\max\limits_{[0,T]}|\gamma|}\,d\boldsymbol{\eta} = 0 \tag{12}$$

for any $t\in[0,T]$. The technical difficulty is that this test function, due to the lack of regularity of $\boldsymbol{b}$, is not continuous. To this aim, we prove first that

$$\int_{\mathbb{R}^d\times\Gamma_T}\frac{\left|\gamma(t)-x-\int_0^t \boldsymbol{c}_s(\gamma(s))\,ds\right|}{1+\max\limits_{[0,T]}|\gamma|}\,d\boldsymbol{\eta} \le \int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s-\boldsymbol{c}_s|}{1+|x|}\,d\mu_s ds \tag{13}$$

for any continuous function $\boldsymbol{c}$ with compact support. Then, choosing a sequence $(\boldsymbol{c}^n)$ converging to $\boldsymbol{b}$ in $L^1(\nu;\mathbb{R}^d)$, with

$$\int \varphi(s,x)\,d\nu(s,x) := \int_0^T\int_{\mathbb{R}^d}\frac{\varphi(s,x)}{1+|x|}\,d\mu_s(x)\,ds$$

and noticing that

$$\int_{\mathbb{R}^d\times\Gamma_T}\int_0^T\frac{|\boldsymbol{b}_s(\gamma(s))-\boldsymbol{c}_s^n(\gamma(s))|}{1+|\gamma(s)|}\,ds d\boldsymbol{\eta} = \int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s-\boldsymbol{c}_s^n|}{1+|x|}\,d\mu_s ds \to 0,$$

we can pass to the limit in (13) with $\boldsymbol{c}=\boldsymbol{c}^n$ to obtain (12).

It remains to show (13). This is a limiting argument based on the fact that (12) holds for $\boldsymbol{b}^\epsilon$, $\boldsymbol{\eta}^\epsilon$:

$$
\begin{aligned}
&\int_{\mathbb{R}^d\times\Gamma_T} \frac{\left|\gamma(t)-x-\int_0^t \boldsymbol{c}_s(\gamma(s))\,ds\right|}{1+\max\limits_{[0,T]}|\gamma|}\,d\boldsymbol{\eta}^\epsilon \\
=\;& \int_{\mathbb{R}^d} \frac{\left|\boldsymbol{X}^\epsilon(t,x)-x-\int_0^t \boldsymbol{c}_s(\boldsymbol{X}^\epsilon(s,x))\,ds\right|}{1+\max\limits_{[0,T]}|\boldsymbol{X}^\epsilon(\cdot,x)|}\,d\mu_0^\epsilon(x) \\
=\;& \int_{\mathbb{R}^d} \frac{\left|\int_0^t \boldsymbol{b}_s^\epsilon(\boldsymbol{X}^\epsilon(s,x))-\boldsymbol{c}_s(\boldsymbol{X}^\epsilon(s,x))\,ds\right|}{1+\max\limits_{[0,T]}|\boldsymbol{X}^\epsilon(\cdot,x)|}\,d\mu_0^\epsilon(x) \le \int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s^\epsilon-\boldsymbol{c}_s|}{1+|x|}\,d\mu_s^\epsilon ds \\
\le\;& \int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s^\epsilon-\boldsymbol{c}_s^\epsilon|}{1+|x|}\,d\mu_s^\epsilon ds+\int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{c}_s^\epsilon-\boldsymbol{c}_s|}{1+|x|}\,d\mu_s^\epsilon ds \\
\le\;& \int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s-\boldsymbol{c}_s|}{1+|x|}\,d\mu_s ds+\int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{c}_s^\epsilon-\boldsymbol{c}_s|}{1+|x|}\,d\mu_s^\epsilon ds.
\end{aligned}
$$

In the last inequalities we added and subtracted $\boldsymbol{c}_t^\epsilon := (\boldsymbol{c}_t\mu_t)*\rho_\epsilon/\mu_t^\epsilon$. Since $\boldsymbol{c}_t^\epsilon \to \boldsymbol{c}_t$ uniformly as $\epsilon\downarrow 0$ thanks to the uniform continuity of $\boldsymbol{c}$, passing to the limit in the chain of inequalities above we obtain (13). $\square$

The applicability of Theorem 9 is strongly limited by the fact that, on one hand, *pointwise* uniqueness properties for the ODE are known only in very special situations, for instance when there is a Lipschitz or a one-sided Lipschitz (or log-Lipschitz, Osgood...) condition on $\boldsymbol{b}$. On the other hand, also uniqueness for general measure-valued solutions is known only in special situations. It turns out that in many cases uniqueness of the PDE can only be proved in smaller classes $\mathscr{L}$ of solutions, and it is natural to think that this should reflect into a weaker uniqueness condition at the level of the ODE.

We will see indeed that there is uniqueness in the "selection sense". In order to illustrate this concept, in the following we consider a convex class $\mathscr{L}_{\boldsymbol{b}}$ of measure-valued solutions $\mu_t\in\mathscr{M}_+(\mathbb{R}^d)$ of the continuity equation relative to $\boldsymbol{b}$, satifying the following monotonicity property:

$$
0\le\mu_t'\le\mu_t\in\mathscr{L}_{\boldsymbol{b}} \qquad\Longrightarrow\qquad \mu_t'\in\mathscr{L}_{\boldsymbol{b}} \tag{14}
$$

whenever μ_t' still solves the continuity equation relative to $\boldsymbol{b}$, and satisfies the integrability condition

$$
\int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_t(x)|}{1+|x|}\,d\mu_t'(x)dt<+\infty.
$$

The typical application will be with absolutely continuous measures $\mu_t = w_t\mathscr{L}^d$, whose densities satisfy some quantitative and possibly time-depending bound (*e.g.* $L^\infty(L^1)\cap L^\infty(L^\infty)$).

Definition 13 ($\mathscr{L}_{\boldsymbol{b}}$-lagrangian flows) *Given the class $\mathscr{L}_{\boldsymbol{b}}$, we say that $\boldsymbol{X}(t,x)$ is a $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar\mu\in\mathscr{M}_+(\mathbb{R}^d)$ (at time 0) if the following two properties hold:*

(a) $\boldsymbol{X}(\cdot, x)$ *is absolutely continuous in* $[0,T]$ *and satisfies*

$$\boldsymbol{X}(t,x) = x + \int_0^t \boldsymbol{b}_s(\boldsymbol{X}(s,x))\,ds \qquad \forall t \in [0,T]$$

for $\bar{\mu}$*-a.e.* x*;*

(b) $\mu_t := \boldsymbol{X}(t,\cdot)_\#\bar{\mu} \in \mathscr{L}_{\boldsymbol{b}}$.

Heuristically $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows can be thought as suitable selections of the (possibly non unique) solutions of the ODE, made in such a way to produce a density in $\mathscr{L}_{\boldsymbol{b}}$. See Example 1 for an illustration of this concept.
We will show that the $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar{\mu}$ is unique, modulo $\bar{\mu}$-negligible sets, whenever uniqueness[1] for the PDE holds in the class $\mathscr{L}_{\boldsymbol{b}}$, *i.e.*

$$\mu_0 = \mu'_0 \qquad \Longrightarrow \qquad \mu_t = \mu'_t \;\; \forall t \in [0,T]$$

whenever μ_t and μ'_t belong to $\mathscr{L}_{\boldsymbol{b}}$.
Before stating and proving the uniqueness theorem for $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows, we state two elementary but useful results. The first one is a simple exercise:

Exercise 14 Let $\sigma \in \mathscr{M}_+(\Gamma_T)$ and let $D \subset [0,T]$ be a dense set. Show that σ is a Dirac mass in Γ_T iff its projections $(e(t))_\#\sigma$, $t \in D$, are Dirac masses in $\mathbb{R}^d$.

The second one is concerned with a family of measures $\boldsymbol{\eta}_x$:

Lemma 15 *Let* $\boldsymbol{\eta}_x$ *be a measurable family of positive finite measures in* Γ_T *with the following property: for any* $t \in [0,T]$ *and any pair of disjoint Borel sets* E, $E' \subset \mathbb{R}^d$ *we have*

$$\boldsymbol{\eta}_x\left(\{\gamma:\ \gamma(t) \in E\}\right)\boldsymbol{\eta}_x\left(\{\gamma:\ \gamma(t) \in E'\}\right) = 0 \quad \bar{\mu}\text{-a.e. in } \mathbb{R}^d. \tag{15}$$

Then $\boldsymbol{\eta}_x$ *is a Dirac mass for* $\bar{\mu}$*-a.e.* x*.*

Proof. Taking into account Exercise 14, for a fixed $t \in (0,T]$ it suffices to check that the measures $\lambda_x := \gamma(t)_\#\boldsymbol{\eta}_x$ are Dirac masses for $\bar{\mu}$-a.e. x. Then (15) gives $\lambda_x(E)\lambda_x(E') = 0$ $\bar{\mu}$-a.e. for any pair of disjoint Borel sets E, $E' \subset \mathbb{R}^d$. Let $\delta > 0$ and let us consider a partition of $\mathbb{R}^d$ in countably many Borel sets R_i having a diameter less then δ. Then, as $\lambda_x(R_i)\lambda_x(R_j) = 0$ $\bar{\mu}$-a.e. whenever $i \neq j$, we have a corresponding decomposition of $\bar{\mu}$-almost all of $\mathbb{R}^d$ in Borel sets A_i such that $\operatorname{supp}\lambda_x \subset \overline{R}_i$ for any $x \in A_i$ (just take $\{\lambda_x(R_i) > 0\}$ and subtract from it all other sets $\{\lambda_x(R_j) > 0\}$, $j \neq i$). Since δ is arbitrary the statement is proved. $\square$

Theorem 16 (Uniqueness of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows) *Assume that the PDE has the uniqueness property in* $\mathscr{L}_{\boldsymbol{b}}$*. Then the* $\mathscr{L}_{\boldsymbol{b}}$*-Lagrangian flow starting from* $\bar{\mu}$ *is unique, i.e. two different*

[1] We thank A. Figalli and F. Flandoli for pointing out that the argument works with minor variants when only uniqueness is imposed at the level of the PDE, and not necessarily the comparison principle

selections $\boldsymbol{X}_1(t,x)$ and $\boldsymbol{X}_2(t,x)$ of solutions of the ODE inducing solutions of the the continuity equation in $\mathscr{L}_{\boldsymbol{b}}$ satisfy

$$\boldsymbol{X}_1(\cdot,x) = \boldsymbol{X}_2(\cdot,x) \qquad \text{in } \Gamma_T, \text{ for } \bar\mu\text{-a.e. } x.$$

More generally, if $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2 \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ are concentrated on the pairs (x,γ) with γ absolutely continuous solution of the ODE, and if $\mu_t^{\boldsymbol{\eta}^1} = \mu_t^{\boldsymbol{\eta}^2} \in \mathscr{L}_{\boldsymbol{b}}$, then $\boldsymbol{\eta}^1 = \boldsymbol{\eta}^2$ and $\boldsymbol{\eta}^i$ are concentrated on the graph of a map $x \mapsto \boldsymbol{X}(\cdot,x)$ which is the unique $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow.

Proof. If the first statement were false we could produce a measure $\boldsymbol{\eta}$ *not* concentrated on a graph inducing a solution $\mu_t^{\boldsymbol{\eta}} \in \mathscr{L}_{\boldsymbol{b}}$ of the PDE. This is not possible, thanks to the next result (Theorem 18). The measure η can be built as follows:

$$\boldsymbol{\eta} := \frac{1}{2}(\boldsymbol{\eta}^1 + \boldsymbol{\eta}^2) = \frac{1}{2}\left[(x, \boldsymbol{X}_1(\cdot,x))_\#\bar\mu + (x, \boldsymbol{X}_2(\cdot,x))_\#\bar\mu\right].$$

Since $\mathscr{L}_{\boldsymbol{b}}$ is convex we still have $\mu_t^{\boldsymbol{\eta}} = \frac{1}{2}(\mu_t^{\boldsymbol{\eta}^1} + \mu_t^{\boldsymbol{\eta}^2}) \in \mathscr{L}_{\boldsymbol{b}}$. In a similar way, one can use Theorem 18 first to show that $\boldsymbol{\eta}^i$ are concentrated on graphs, and then the previous combination argument to show that $\boldsymbol{\eta}^1 = \boldsymbol{\eta}^2$. $\square$

Remark 17 In the same vein, one can also show that

$$\boldsymbol{X}_1(\cdot,x) = \boldsymbol{X}_2(\cdot,x) \qquad \text{in } \Gamma_T \text{ for } \bar\mu_1 \wedge \bar\mu_2\text{-a.e. } x$$

whenever $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ are $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows starting respectively from $\bar\mu_1$ and $\bar\mu_2$.

We used the following basic result, having some analogy with Kantorovich's and Mather's theories.

Theorem 18 *Assume that the PDE has the uniqueness property in $\mathscr{L}_{\boldsymbol{b}}$. Let $\boldsymbol{\eta} \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ be concentrated on the pairs (x,γ) with γ absolutely continuous solution of the ODE starting from x, and assume that $\mu_t^{\boldsymbol{\eta}} \in \mathscr{L}_{\boldsymbol{b}}$. Then $\boldsymbol{\eta}$ is concentrated on a graph, i.e. there exists a function $x \mapsto X(\cdot,x) \in \Gamma_T$ such that*

$$\boldsymbol{\eta} = \big(x, X(\cdot,x)\big)_\#\bar\mu, \quad \text{with} \quad \bar\mu := (\pi_{\mathbb{R}^d})_\#\boldsymbol{\eta} = \mu_0^{\boldsymbol{\eta}}.$$

Proof. We use the representation (8) of $\boldsymbol{\eta}$, given by the disintegration theorem, the criterion stated in Lemma 15 and argue by contradiction. If the thesis is false then $\boldsymbol{\eta}_x$ is not a Dirac mass in a set of $\bar\mu$ positive measure and we can find $t \in (0,T]$, disjoint Borel sets E, $E' \subset \mathbb{R}^d$ and a Borel set C with $\bar\mu(C) > 0$ such that

$$\boldsymbol{\eta}_x\big(\{\gamma : \ \gamma(t) \in E\}\big)\boldsymbol{\eta}_x\big(\{\gamma : \ \gamma(t) \in E'\}\big) > 0 \qquad \forall x \in C.$$

Possibly passing to a smaller set having still strictly positive $\bar\mu$ measure we can assume that

$$0 < \boldsymbol{\eta}_x(\{\gamma : \ \gamma(t) \in E\}) \le M\boldsymbol{\eta}_x(\{\gamma : \ \gamma(t) \in E'\}) \qquad \forall x \in C \tag{16}$$

for some constant M. We define measures $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2$ whose disintegrations $\boldsymbol{\eta}^1_x$, $\boldsymbol{\eta}^2_x$ are given by

$$\boldsymbol{\eta}^1_x := \chi_C(x)\boldsymbol{\eta}_x \llcorner \{\gamma: \ \gamma(t) \in E\}, \qquad \boldsymbol{\eta}^2_x := M\chi_C(x)\boldsymbol{\eta}_x \llcorner \{\gamma: \ \gamma(t) \in E'\}$$

and denote by μ^i_s, $s \in [0,t]$, the (superposition) solutions of the continuity equation induced by $\boldsymbol{\eta}^i$. Then

$$\mu^1_0 = \boldsymbol{\eta}_x(\{\gamma: \ \gamma(t) \in E\})\bar{\mu}\llcorner C, \qquad \mu^2_0 = M\boldsymbol{\eta}_x(\{\gamma: \ \gamma(t) \in E'\})\bar{\mu}\llcorner C,$$

so that (16) yields $\mu^1_0 \leq \mu^2_0$. On the other hand, μ^1_t is orthogonal to μ^2_t: precisely, denoting by $\boldsymbol{\eta}_{tx}$ the image of $\boldsymbol{\eta}_x$ under the map $\gamma \mapsto \gamma(t)$, we have

$$\mu^1_t = \int_C \boldsymbol{\eta}_{tx} \llcorner E \, d\mu(x) \perp M \int_C \boldsymbol{\eta}_{tx} \llcorner E' \, d\mu(x) = \mu^2_t.$$

In order to conclude, let $f : \mathbb{R}^d \to [0,1]$ be the density of μ^1_0 with respect to μ^2_0 and set

$$\tilde{\boldsymbol{\eta}}^2_x := Mf(x)\chi_C(x)\boldsymbol{\eta}_x \llcorner \{\gamma: \ \gamma(t) \in E'\}.$$

We define the measure $\tilde{\boldsymbol{\eta}}^2$ whose disintegration is given by $\tilde{\boldsymbol{\eta}}^2_x$ and denote by $\tilde{\mu}^2_s$, $s \in [0,t]$, the (superposition) solution of the continuity equation induced by $\tilde{\boldsymbol{\eta}}^2$.
Notice also that $\mu^i_s \leq \mu_s$ and so the monotonicity assumption (14) on $\mathscr{L}_{\boldsymbol{b}}$ gives $\mu^i_s \in \mathscr{L}_{\boldsymbol{b}}$, and since $\tilde{\boldsymbol{\eta}}^2 \leq \boldsymbol{\eta}^2$ we obtain that $\tilde{\mu}^2_s \in \mathscr{L}_{\boldsymbol{b}}$ as well. By construction $\mu^1_0 = \tilde{\mu}^2_0$, while μ^1_t is orthogonal to μ^2_t, a measure larger than $\tilde{\mu}^2_t$. We have thus built two different solutions of the PDE with the same initial condition. □

Now we come to the *existence* of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows.

Theorem 19 (Existence of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows) *Assume that the PDE has the uniqueness property in $\mathscr{L}_{\boldsymbol{b}}$ and that for some $\bar{\mu} \in \mathscr{M}_+(\mathbb{R}^d)$ there exists a solution $\mu_t \in \mathscr{L}_{\boldsymbol{b}}$ with $\mu_0 = \bar{\mu}$. Then there exists a (unique) $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar{\mu}$.*

Proof. By the superposition principle we can represent μ_t as $(e_t)_\#\boldsymbol{\eta}$ for some $\boldsymbol{\eta} \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ concentrated on pairs (x,γ) solutions of the ODE. Then, Theorem 18 tells us that $\boldsymbol{\eta}$ is concentrated on a graph, i.e. there exists a function $x \mapsto \boldsymbol{X}(\cdot,x) \in \Gamma_T$ such that

$$\big(x, \boldsymbol{X}(\cdot,x)\big)_\#\bar{\mu} = \boldsymbol{\eta}.$$

Pushing both sides via e_t we obtain

$$\boldsymbol{X}(t,\cdot)_\#\bar{\mu} = (e_t)_\#\boldsymbol{\eta} = \mu_t \in \mathscr{L}_{\boldsymbol{b}},$$

and therefore $\boldsymbol{X}$ is a $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow. □

5 The flow associated to Sobolev or BV vector fields

Here we discuss the well-posedness of the continuity or transport equations assuming that $\boldsymbol{b}_t(\cdot)$ has a Sobolev regularity, following [61]. Then, the general theory previously developed provides existence and uniqueness of the $\mathscr{L}$-Lagrangian flow, with $\mathscr{L} := L^\infty(L^1) \cap L^\infty(L^\infty)$. We denote by $I \subset \mathbb{R}$ an open interval.

Definition 20 (Renormalized solutions) *Let $\boldsymbol{b} \in L^1_{\rm loc}\left(I; L^1_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ be such that $D \cdot \boldsymbol{b}_t = \operatorname{div} \boldsymbol{b}_t \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in I$, with*

$$\operatorname{div} \boldsymbol{b}_t \in L^1_{\rm loc}\Big(I; L^1_{\rm loc}(\mathbb{R}^d)\Big).$$

Let $w \in L^\infty_{\rm loc}\left(I; L^\infty_{\rm loc}(\mathbb{R}^d)\right)$ and assume that

$$c := \frac{d}{dt} w + \boldsymbol{b} \cdot \nabla w \in L^1_{\rm loc}(I \times \mathbb{R}^d). \tag{17}$$

Then, we say that w is a renormalized solution of (17) *if*

$$\frac{d}{dt}\beta(w) + \boldsymbol{b} \cdot \nabla \beta(w) = c\beta'(w) \qquad \forall \beta \in C^1(\mathbb{R}).$$

Equivalently, recalling the definition of the distribution $\boldsymbol{b} \cdot \nabla w$, the definition could be given in a conservative form, writing

$$\frac{d}{dt}\beta(w) + D_x \cdot (\boldsymbol{b}\beta(w)) = c\beta'(w) + \operatorname{div} \boldsymbol{b}_t \beta(w).$$

Notice also that the concept makes sense, choosing properly the class of "test" functions β, also for w not satisfying (17), or not even locally integrable. This is particularly relevant in connection with DiPerna-Lions's existence theorem for Boltzmann equation [62], or with the case when w is the characteristic of an unbounded vector field $\boldsymbol{b}$.
This concept is also reminiscent of Kruzkhov's concept of *entropy* solution for a scalar conservation law

$$\frac{d}{dt}u + D_x \cdot (\boldsymbol{f}(u)) = 0 \qquad u : (0, +\infty) \times \mathbb{R}^d \to \mathbb{R}.$$

In this case only a distributional one-sided inequality is required:

$$\frac{d}{dt}\eta(u) + D_x \cdot (\boldsymbol{q}(u)) \le 0$$

for any convex entropy-entropy flux pair $(\eta, \boldsymbol{q})$ (i.e. η is convex and $\eta' \boldsymbol{f}' = \boldsymbol{q}'$).

Remark 21 (Time continuity) Using the fact that both $t \mapsto w_t$ and $t \mapsto \beta(w_t)$ have uniformly continuous representatives (w.r.t. the $w^* - L^\infty_{\rm loc}$ topology) $\bar{w}_t$, σ_t, we obtain that $t \mapsto \bar{w}_t$ is continuous with respect to the strong $L^1_{\rm loc}$ topology for $\mathscr{L}^1$-a.e. t, and precisely for any t such that $\sigma_t = \beta(\bar{w}_t)$. The proof follows by a classical weak-strong convergence argument:

$$f_n \rightharpoonup f, \quad \beta(f_n) \rightharpoonup \beta(f) \qquad \Longrightarrow \qquad f_n \to f$$

provided β is *strictly* convex. In the case of scalar conservation laws there are analogous and more precise results [92], [82]. We remark the fact that, in general, a renormalized solution does *not* need to have a representative which is strongly continuous *for every* t. This can be seen using a variation of an example given by Depauw [59] (see Remark 2.7 of [26]). Depauw's example provides a divergence free vector field $\boldsymbol{a} \in L^\infty([0,1] \times \mathbb{R}^2; \mathbb{R}^2)$, with $\boldsymbol{a}(t,\cdot) \in BV_{\mathrm{loc}}(\mathbb{R}^2;\mathbb{R}^2)$ for $\mathscr{L}^1$-a.e. $t \in [0,1]$ (but $\boldsymbol{a} \notin L^1([0,1]; BV_{\mathrm{loc}})$) such that the Cauchy problem

$$\begin{cases} \partial_t u + \boldsymbol{a} \cdot \nabla u = 0 \\ u(0,\cdot) = 0 \end{cases}$$

has a nontrivial solution, with $|\bar{u}| = 1$ $\mathscr{L}^3$-a.e. in $[0,1] \times \mathbb{R}^2$ and with the property that $\bar{u}(t,\cdot) \rightharpoonup 0$ as $t \downarrow 0$, but this convergence is *not* strong. Now consider a vector field $\boldsymbol{b}$ on $[-1,1] \times \mathbb{R}^2$ defined as Depauw's vector field for $t > 0$, and set $\boldsymbol{b}(t,x) = -\boldsymbol{a}(-t,x)$ for $t < 0$. It is simple to check (as only affine functions $\tilde{\beta}(t) = a + bt$ need to be checked, because for any β there exists an affine $\tilde{\beta}$ such that $\tilde{\beta}(\pm 1) = \beta(\pm 1)$) that the function

$$\bar{w}(t,x) = \begin{cases} \bar{u}(t,x) & \text{if } t > 0 \\ \bar{u}(-t,x) & \text{if } t < 0 \end{cases}$$

is a renormalized solution of $\partial_t w + \boldsymbol{b} \cdot \nabla w = 0$, but this solution is not strongly continuous at $t = 0$.

Remark 22 A new insight in the theory of renormalized solutions has been obtained in the recent paper [26]. In particular, it is proved that for a vector field $\boldsymbol{b} \in L^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ with zero divergence (and without any regularity assumption) the following two conditions are equivalent (the L^2 framework has been considered just for simplicity):

(i) $\boldsymbol{b}$ has the uniqueness property for weak solutions in $C([0,T]; w - L^2(\mathbb{R}^d))$ for both the forward and the backward Cauchy problems starting respectively from 0 and T, i.e. the only solutions in $C([0,T]; w - L^2(\mathbb{R}^d))$ to the problems

$$\begin{cases} \partial_t u_F + \boldsymbol{b} \cdot \nabla u_F = 0 \\ u_F(0,\cdot) = 0 \end{cases} \quad \text{and} \quad \begin{cases} \partial_t u_B + \boldsymbol{b} \cdot \nabla u_B = 0 \\ u_B(T,\cdot) = 0 \end{cases}$$

are $u_F \equiv 0$ and $u_B \equiv 0$;

(ii) every weak solution in $C([0,T]; w - L^2(\mathbb{R}^d))$ of $\partial_t u + \boldsymbol{b} \cdot \nabla u = 0$ is strongly continuous (i.e. lies in $C([0,T]; s - L^2(\mathbb{R}^d))$) and is a renormalized solution.

The proof of this equivalence is obtained through the study of the approximation properties of the solution of the transport equation, with respect to the norm of the graph of the transport operator (see Theorem 2.1 of [26] for the details).

Using the concept of renormalized solution we can prove a comparison principle in the following natural class $\mathscr{L}$:

$$\begin{aligned} \mathscr{L} \quad := \quad & \Big\{ w \in L^\infty\Big([0,T]; L^1(\mathbb{R}^d)\Big) \cap L^\infty\Big([0,T]; L^\infty(\mathbb{R}^d)\Big) : \\ & w \in C\Big([0,T]; w^* - L^\infty(\mathbb{R}^d)\Big)\Big\}. \end{aligned} \tag{18}$$

Theorem 23 (Comparison principle) *Assume that*

$$\frac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T];L^\infty(\mathbb{R}^d)\right) + L^1\left([0,T];L^1(\mathbb{R}^d)\right), \tag{19}$$

that $D\cdot\boldsymbol{b}_t = \operatorname{div}\boldsymbol{b}_t\mathscr{L}^d$ *for* $\mathscr{L}^1$*-a.e.* $t\in[0,T]$*, and that*

$$[\operatorname{div}\boldsymbol{b}_t]^- \in L^1_{\rm loc}\left([0,T)\times\mathbb{R}^d\right). \tag{20}$$

Setting $\boldsymbol{b}_t\equiv 0$ *for* $t<0$*, assume in addition that* any *solution of* (17) *in* $(-\infty,T)\times\mathbb{R}^d$ *is renormalized. Then the comparison principle for the continuity equation holds in the class* $\mathscr{L}$*.*

Proof. By the linearity of the equation, it suffices to show that $w\in\mathscr{L}$ and $w_0\le 0$ implies $w_t\le 0$ for any $t\in[0,T]$. We extend first the PDE to negative times, setting $w_t=w_0$. Then, fix a cut-off function $\varphi\in C^\infty_c(\mathbb{R}^d)$ with $\operatorname{supp}\varphi\subset\overline{B}_2(0)$ and $\varphi\equiv 1$ on $B_1(0)$, and the renormalization functions

$$\beta_\epsilon(t) := \sqrt{\epsilon^2+(t^+)^2}-\epsilon\in C^1(\mathbb{R}).$$

Notice that

$$\beta_\epsilon(t)\uparrow t^+ \quad\text{as } \epsilon\downarrow 0, \qquad t\beta_\epsilon'(t)-\beta_\epsilon(t)\in[0,\epsilon]. \tag{21}$$

We know that

$$\frac{d}{dt}\beta_\epsilon(w_t)+D_x\cdot(\boldsymbol{b}\beta_\epsilon(w_t)) = \operatorname{div}\boldsymbol{b}_t(\beta_\epsilon(w_t)-w_t\beta_\epsilon'(w_t))$$

in the sense of distributions in $(-\infty,T)\times\mathbb{R}^d$. Plugging $\varphi_R(\cdot):=\varphi(\cdot/R)$, with $R\ge 1$, into the PDE we obtain

$$\frac{d}{dt}\int_{\mathbb{R}^d}\varphi_R\beta_\epsilon(w_t)\,dx = \int_{\mathbb{R}^d}\beta_\epsilon(w_t)\langle\boldsymbol{b}_t,\nabla\varphi_R\rangle\,dx + \int_{\mathbb{R}^d}\varphi_R\operatorname{div}\boldsymbol{b}_t(\beta_\epsilon(w_t)-w_t\beta_\epsilon'(w_t))\,dx.$$

Splitting $\boldsymbol{b}$ as $\boldsymbol{b}_1+\boldsymbol{b}_2$, with

$$\frac{|\boldsymbol{b}_1|}{1+|x|}\in L^1\left([0,T];L^\infty(\mathbb{R}^d)\right) \quad\text{and}\quad \frac{|\boldsymbol{b}_2|}{1+|x|}\in L^1\left([0,T];L^1(\mathbb{R}^d)\right)$$

and using the inequality

$$\frac{1}{R}\chi_{\{R\le|x|\le 2R\}} \le \frac{3}{1+|x|}\chi_{\{R\le|x|\}}$$

we can estimate the first integral in the right hand side with

$$3\|\nabla\varphi\|_\infty\left\|\frac{\boldsymbol{b}_{1t}}{1+|x|}\right\|_\infty\int_{\{|x|\ge R\}}|w_t|\,dx + 3\|\nabla\varphi\|_\infty\|w_t\|_\infty\int_{\{|x|\ge R\}}\frac{|\boldsymbol{b}_{1t}|}{1+|x|}\,dx.$$

The second integral can be estimated with

$$\epsilon\int_{\mathbb{R}^d}\varphi_R[\operatorname{div}\boldsymbol{b}_t]^-\,dx.$$

Passing to the limit first as $\epsilon \downarrow 0$ and then as $R \to +\infty$ and using the integrability assumptions on $\boldsymbol{b}$ and w we get

$$\frac{d}{dt}\int_{\mathbb{R}^d} w_t^+ \, dx \le 0$$

in the distribution sense in $\mathbb{R}$. Since the function vanishes for negative times, this suffices to conclude using Gronwall lemma. $\square$

DiPerna and Lions proved that *all* distributional solutions are renormalized when there is a Sobolev regularity with respect to the spatial variables.

Theorem 24 *Let $\boldsymbol{b} \in L^1_{\rm loc}\left(I; W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ and let $w \in L^\infty_{\rm loc}(I\times\mathbb{R}^d)$ be a distributional solution of* (17)*. Then w is a renormalized solution.*

Proof. We mollify with respect to the spatial variables and we set

$$r^\epsilon := (\boldsymbol{b}\cdot\nabla w) * \rho_\epsilon - \boldsymbol{b}\cdot(\nabla(w*\rho_\epsilon)), \qquad w^\epsilon := w*\rho_\epsilon$$

to obtain

$$\frac{d}{dt}w^\epsilon + \boldsymbol{b}\cdot\nabla w^\epsilon = c*\rho_\epsilon - r^\epsilon.$$

By the smoothness of w^ϵ w.r.t. x, the PDE above tells us that $\frac{d}{dt}w^\epsilon_t \in L^1_{\rm loc}$, therefore $w^\epsilon \in W^{1,1}_{\rm loc}(I\times\mathbb{R}^d)$ and we can apply the standard chain rule in Sobolev spaces, getting

$$\frac{d}{dt}\beta(w^\epsilon) + \boldsymbol{b}\cdot\nabla\beta(w^\epsilon) = \beta'(w^\epsilon)c*\rho_\epsilon - \beta'(w^\epsilon)r^\epsilon.$$

When we let $\epsilon \downarrow 0$, the convergence in the distribution sense of all terms in the identity above is trivial, with the exception of the last one. To ensure its convergence to zero, it seems necessary to show that $r^\epsilon \to 0$ strongly in $L^1_{\rm loc}$, or at least that r^ϵ are equi-integrable (as this would imply their weak L^1 convergence to 0; remember also that $\beta'(w^\epsilon)$ is locally equibounded w.r.t. ϵ). This is indeed the case, and it is exactly here that the Sobolev regularity plays a role. $\square$

Proposition 25 (Strong convergence of commutators) *If $w \in L^\infty_{\rm loc}(I\times\mathbb{R}^d)$ and $\boldsymbol{b} \in L^1_{\rm loc}\left(I; W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ we have*

$$L^1_{\rm loc}\text{-}\lim_{\epsilon\downarrow 0}\,(\boldsymbol{b}\cdot\nabla w)*\rho_\epsilon - \boldsymbol{b}\cdot(\nabla(w*\rho_\epsilon)) = 0.$$

Proof. Playing with the definitions of $\boldsymbol{b}\cdot\nabla w$ and of the convolution product of a distribution and a smooth function, one proves first the identity

$$r^\epsilon(t,x) = \int_{\mathbb{R}^d} w(t,x-\epsilon y)\frac{(\boldsymbol{b}_t(x-\epsilon y)-\boldsymbol{b}_t(x))\cdot\nabla\rho(y)}{\epsilon}\,dy - (w{\rm div}\,\boldsymbol{b}_t)*\rho_\epsilon(x). \tag{22}$$

Introducing the commutators in the (easier) conservative form

$$R^\epsilon := (D_x\cdot(\boldsymbol{b}w))*\rho_\epsilon - D_x\cdot(\boldsymbol{b}w^\epsilon)$$

(here we set again $w^\epsilon := w * \rho_\epsilon$) it suffices to show that $R^\epsilon = L^\epsilon - w^\epsilon \mathrm{div}\, \boldsymbol{b}_t$, where

$$L^\epsilon(t,x) := \int_{\mathbb{R}^d} w(t,z)(\boldsymbol{b}_t(x) - \boldsymbol{b}_t(z)) \cdot \nabla \rho_\epsilon(z-x)\, dz.$$

Indeed, for any test function φ, we have that $\langle R^\epsilon, \varphi\rangle$ is given by

$$\begin{aligned}
& -\int_I \int w\boldsymbol{b}\cdot\nabla\rho_\epsilon * \varphi dydt - \int_I\int \varphi\boldsymbol{b}\cdot\nabla\rho_\epsilon * w dxdt - \int_I\int w^\epsilon\varphi \mathrm{div}\,\boldsymbol{b}_t dxdt \\
= & -\int_I\int\int w_t(y)\boldsymbol{b}_t(y)\cdot\nabla\rho_\epsilon(y-x)\varphi(x)\, dxdydt \\
& -\int_I\int\int \boldsymbol{b}_t(x)\nabla\rho_\epsilon(x-y)w_t(y)\varphi(x) dydxdt - \int_I\int w^\epsilon\varphi\mathrm{div}\,\boldsymbol{b}_t\, dxdt \\
= & \int_I\int L^\epsilon\varphi\, dxdt - \int_I\int w^\epsilon\varphi\mathrm{div}\,\boldsymbol{b}_t\, dxdt
\end{aligned}$$

(in the last equality we used the fact that $\nabla\rho$ is odd).
Then, one uses the strong convergence of translations in L^p and the strong convergence of the difference quotients (a property that *characterizes* functions in Sobolev spaces)

$$\frac{u(x+\epsilon z) - u(x)}{\epsilon} \to \nabla u(x) z \qquad \text{strongly in } L^1_{\mathrm{loc}}, \text{ for } u \in W^{1,1}_{\mathrm{loc}}$$

to obtain that r^ϵ strongly converge in $L^1_{\mathrm{loc}}(I\times\mathbb{R}^d)$ to

$$-w(t,x)\int_{\mathbb{R}^d}\langle\nabla\boldsymbol{b}_t(x)y, \nabla\rho(y)\rangle\, dy - w(t,x)\mathrm{div}\,\boldsymbol{b}_t(x).$$

The elementary identity

$$\int_{\mathbb{R}^d} y_i \frac{\partial\rho}{\partial y_j}\, dy = -\delta_{ij}$$

then shows that the limit is 0 (this can also be derived by the fact that, in any case, the limit of r^ϵ in the distribution sense should be 0). $\square$

In this context, given $\bar\mu = \rho\mathscr{L}^d$ with $\rho \in L^1\cap L^\infty$, the $\mathscr{L}$-Lagrangian flow starting from $\bar\mu$ (at time 0) is defined by the following two properties:

(a) $\boldsymbol{X}(\cdot, x)$ is absolutely continuous in $[0,T]$ and satisfies

$$\boldsymbol{X}(t,x) = x + \int_0^t \boldsymbol{b}_s(\boldsymbol{X}(s,x))\, ds \qquad \forall t\in[0,T]$$

for $\bar\mu$-a.e. x;

(b) $\boldsymbol{X}(t,\cdot)_{\#}\bar\mu \le C\mathscr{L}^d$ for all $t\in[0,T]$, with C independent of t.

Summing up what we obtained so far, the general theory provides us with the following existence and uniqueness result.

Theorem 26 (Existence and uniqueness of $\mathscr{L}$-Lagrangian flows) *Let* $\boldsymbol{b} \in L^1\left([0,T]; W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ *be satisfying*

(i) $\dfrac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T];L^1(\mathbb{R}^d)\right) + L^1\left([0,T];L^\infty(\mathbb{R}^d)\right)$*;*

(ii) $[\operatorname{div}\boldsymbol{b}_t]^- \in L^1\left([0,T];L^\infty(\mathbb{R}^d)\right)$.

Then the $\mathscr{L}$-Lagrangian flow relative to $\boldsymbol{b}$ exists and is unique.

Proof. By the previous results, the comparison principle holds for the continuity equation relative to $\boldsymbol{b}$. Therefore the general theory previously developed applies, and Theorem 16 provides *uniqueness* of the $\mathscr{L}$-Lagrangian flow.
As for the *existence*, still the general theory (Theorem 19) tells us that it can be achieved provided we are able to solve, within $\mathscr{L}$, the continuity equation

$$\frac{d}{dt}w + D_x\cdot(\boldsymbol{b}w) = 0 \tag{23}$$

for any nonnegative initial datum $w_0 \in L^1\cap L^\infty$. The existence of these solutions can be immediately achieved by a smoothing argument: we approximate $\boldsymbol{b}$ in $L^1_{\rm loc}$ by smooth $\boldsymbol{b}^h$ with a uniform bound in $L^1(L^\infty)$ for $[\operatorname{div}\boldsymbol{b}^h_t]^-$. This bound, in turn, provides a uniform lower bound on $J\boldsymbol{X}^h$ and finally a uniform upper bound on $w^h_t = (w_0/J\boldsymbol{X}^h_t)\circ(\boldsymbol{X}^h_t)^{-1}$, solving

$$\frac{d}{dt}w^h + D_x\cdot(\boldsymbol{b}^h w^h) = 0.$$

Therefore, any weak limit of w^h solves (23). □

Notice also that, choosing for instance a Gaussian, we obtain that the $\mathscr{L}$-Lagrangian flow is well defined up to $\mathscr{L}^d$-negligible sets (and independent of $\bar\mu \ll \mathscr{L}^d$, thanks to Remark 17).
It is interesting to compare our characterization of Lagrangian flows with the one given in [61]. Heuristically, while the DiPerna-Lions characterization is based on the semigroup of transformations $x\mapsto \boldsymbol{X}(t,x)$, our characterization is based on the properties of the map $x\mapsto\boldsymbol{X}(\cdot,x)$.

Remark 27 The definition of the flow in [61] is based on the following three properties:

(a) $\dfrac{\partial\boldsymbol{Y}}{\partial t}(t,s,x) = b\left(t,\boldsymbol{Y}(t,s,x)\right)$ and $\boldsymbol{Y}(s,s,x)=x$ in the distribution sense in $(0,T)\times\mathbb{R}^d$;

(b) the image λ_t of $\mathscr{L}^d$ under $\boldsymbol{Y}(t,s,\cdot)$ satisfies

$$\frac{1}{C}\mathscr{L}^d \le \lambda_t \le C\mathscr{L}^d \qquad \text{for some constant } C>0;$$

(c) for all $s,\ s',t\in[0,T]$ we have

$$\boldsymbol{Y}\left(t,s,\boldsymbol{Y}(s,s',x)\right) = \boldsymbol{Y}(t,s',x) \qquad \text{for } \mathscr{L}^d\text{-a.e. } x.$$

Then, $\boldsymbol{Y}(t,s,x)$ corresponds, in our notation, to the flow $\boldsymbol{X}^s(t,x)$ starting at time s (well defined even for $t<s$ if one has two-sided L^∞ bounds on the divergence).
In our setting condition (c) can be recovered as a consequence of the following argument: assume to fix the ideas that $s'\le s\le T$ and define

$$\tilde{\boldsymbol{X}}(t,x) := \begin{cases} \boldsymbol{X}^{s'}(t,x) & \text{if } t\in[s',s]; \\ \\ \boldsymbol{X}^{s}\Big(t,\boldsymbol{X}^{s'}(s,x)\Big) & \text{if } t\in[s,T]. \end{cases}$$

It is immediate to check that $\tilde{\boldsymbol{X}}(\cdot,x)$ is an integral solution of the ODE in $[s',T]$ for $\mathscr{L}^d$-a.e. x and that $\tilde{\boldsymbol{X}}(t,\cdot)_\#\bar\mu$ is bounded by $C^2\mathscr{L}^d$. Then, Theorem 26 (with s' as initial time) gives $\tilde{\boldsymbol{X}}(\cdot,x)=\boldsymbol{X}(\cdot,s',x)$ in $[s',T]$ for $\mathscr{L}^d$-a.e. x, whence (c) follows.

Let us now discuss the stability properties of $\mathscr{L}$-Lagrangian flows, in the special case when $\mathscr{L}$ is defined as in (18). We need the following lemma.

Lemma 28 *Assume that $\boldsymbol{b}_h:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ satisfy*

$$\frac{\boldsymbol{b}_h}{1+|x|}\in L^1\Big([0,T];L^1(\mathbb{R}^d;\mathbb{R}^d)\Big)+L^1\Big([0,T];L^\infty(\mathbb{R}^d;\mathbb{R}^d)\Big) \tag{24}$$

and that we can write $\boldsymbol{b}_h=\boldsymbol{b}_h^1+\boldsymbol{b}_h^2$, with

$$\frac{|\boldsymbol{b}_h^1|}{1+|x|}\quad \textit{bounded and equi-integrable in } L^1\Big([0,T]\times\mathbb{R}^d\Big), \tag{25}$$

$$\sup_h\left\|\frac{|\boldsymbol{b}_h^2(t,\cdot)|}{1+|x|}\right\|_\infty\in L^1(0,T), \tag{26}$$

$$\boldsymbol{b}_h^1\to\boldsymbol{b}^1,\quad \boldsymbol{b}_h^2\to\boldsymbol{b}^2\qquad \mathscr{L}^{d+1}\textit{-a.e. in } (0,T)\times\mathbb{R}^d. \tag{27}$$

Then, setting $\boldsymbol{b}=\boldsymbol{b}^1+\boldsymbol{b}^2$, we have

$$\lim_{h\to\infty}\int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_h-\boldsymbol{b}|}{1+|x|^{d+2}}\,dxdt=0. \tag{28}$$

Proof. Without loss of generality we can assume that $\boldsymbol{b}^1=\boldsymbol{b}^2=0$, hence $\boldsymbol{b}=0$. The convergence to 0 of

$$\int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_h^2|}{1+|x|^{d+2}}\,dxdt$$

follows by the standard dominated convergence theorem. The convergence to 0 of

$$\int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_h^1|}{1+|x|}\,dxdt$$

(and a fortiori of the integrals with the factor $1+|x|^{d+2}$) follows by the Vitali dominated convergence theorem. □

Theorem 29 *Let $\boldsymbol{X}_h$ be $\mathscr{L}$-Lagrangian flows relative to vector fields $\boldsymbol{b}_h$, starting from $\bar\mu = \bar\rho\mathscr{L}^{2d}$ and satisfying:*

(a) $\boldsymbol{X}_h(t,\cdot)_\#\bar\mu \le C\mathscr{L}^d$, with C independent of h and $t \in [0,T]$;

(b) $\boldsymbol{b}_h = \boldsymbol{b}_h^1 + \boldsymbol{b}_h^2$ with $\boldsymbol{b}_h^i$ satisfying (25), (26) *and* (27).

Assume that the continuity equation with $\boldsymbol{b} = \boldsymbol{b}^1 + \boldsymbol{b}^2$ as vector field satisfies the uniqueness property in $\mathscr{L}$. Then there exists a unique $\mathscr{L}$-Lagrangian flow $\boldsymbol{X}$ relative to $\boldsymbol{b}$ and $x \mapsto \boldsymbol{X}^h(\cdot,x)$ converge to $x \mapsto \boldsymbol{X}(\cdot,x)$ in $\bar\mu$-measure, i.e.

$$\lim_{h\to\infty} \int_{\mathbb{R}^d} 1 \wedge \max_{[0,T]} |\boldsymbol{X}^h(t,x) - \boldsymbol{X}(t,x)|\, d\bar\mu(x) = 0.$$

Remark 30 (Stability with weak convergence in time) We remark the fact that the hypothesis of strong convergence in both time and space of the vector fields in the stability theorem is not natural in view of the applications to the theory of fluid mechanics (see Theorem II.7 in [61] and [75], in particular Theorem 2.5), and it is in contrast with the general philosophy that time regularity is less important than spatial regolarity (only summability with respect to time is necessary for the renormalization property to hold). For instance, this form of the stability theorem does not include the case of weakly converging vector fields which depend *only* on the time variable, while it is clear that in this case the convergence of the flows holds. However, as a consequence of the quantitative estimates presented in Section 8, it is possible to show that, under uniform bounds in $L^\infty([0,T];W^{1,p}_{\rm loc})$ for some $p > 1$ (and for simplicity under uniform bounds in L^∞), the following form of weak convergence with respect to the time is sufficient to get the thesis:

$$\int_0^T \boldsymbol{b}_h(t,x)\eta(t)\,dt \longrightarrow \int_0^T \boldsymbol{b}(t,x)\eta(t)\,dt \qquad \text{in } L^1_{\rm loc}(\mathbb{R}^d) \text{ for every } \eta \in C^\infty_c(0,T).$$

Indeed, fix a parameter $\epsilon > 0$ and regularize with respect to the spatial variables only using a standard convolution kernel ρ_ϵ. We can rewrite the difference $\boldsymbol{X}_h(t,x) - \boldsymbol{X}(t,x)$ as

$$\boldsymbol{X}_h(t,x) - \boldsymbol{X}(t,x) = \Big(\boldsymbol{X}_h(t,x) - \boldsymbol{X}_h^\epsilon(t,x)\Big) + \Big(\boldsymbol{X}_h^\epsilon(t,x) - \boldsymbol{X}^\epsilon(t,x)\Big) + \Big(\boldsymbol{X}^\epsilon(t,x) - \boldsymbol{X}(t,x)\Big),$$

where $\boldsymbol{X}^\epsilon$ and $\boldsymbol{X}_h^\epsilon$ are the flows relative to the spatial regularizations $\boldsymbol{b}^\epsilon$ and $\boldsymbol{b}_h^\epsilon$ respectively. Now, it is simple to check that

- the last term goes to zero with ϵ, by the stability theorem stated above;
- the first term goes to zero with ϵ, uniformly with respect to h: this is due to the fact that the difference $\boldsymbol{b}_h^\epsilon - \boldsymbol{b}_h$ goes to zero in $L^1_{\rm loc}([0,T]\times\mathbb{R}^d)$ uniformly with respect to h (thanks to the uniform control in $W^{1,p}$ of the vector fields $\boldsymbol{b}_h$), hence we can apply the quantitative version of the stability theorem (see Theorem 53), and we get the desired convergence;

- the second term goes to zero for $h \to \infty$ when ϵ is kept fixed, because we are dealing with flows relative to vector fields which are smooth with respect to the space variable, uniformly in time, and weak convergence with respect to the time is enough to get the stability.

Then, in order to conclude, it is enough to let first $h \to \infty$, to eliminate the second term, and then $\epsilon \to 0$.
It is not clear to us whether this result extends to the $W^{1,1}$ or the BV case: the quantitative estimates are available only for $W^{1,p}$ vector fields, where p is strictly greater than 1, hence the strategy described before does not extend to that case.

Proof. (*of Theorem 29*) We define $\boldsymbol{\eta}_h$ as the push forward of $\bar{\mu}$ under the map $x \mapsto (x, \boldsymbol{X}^h(\cdot, x))$ and argue as in the proof of Theorem 12.
Step 1. (Tightness of $\boldsymbol{\eta}_h$) We claim that the family $\boldsymbol{\eta}_h$ is tight: indeed, by the remarks made after the statement of Theorem 12, it suffices to find a coercive functional $\Psi : \mathbb{R}^d \times \Gamma_T \to [0, \infty)$ whose integrals w.r.t. all measures $\boldsymbol{\eta}_h$ are uniformly bounded. Since $\bar{\mu}$ has finite mass we can find a function $\varphi : \mathbb{R}^d \to [0, \infty)$ such that $\varphi \in L^1(\bar{\mu})$ and $\varphi(x) \to \infty$ as $|x| \to \infty$. Analogously, we can find a function $\psi : \mathbb{R}^d \to [0, \infty)$ such that

$$\int_0^T \int_{\mathbb{R}^d} \frac{\psi(\boldsymbol{b})}{1 + |x|}\, dxdt < \infty$$

and $\psi(x) \to \infty$ as $|x| \to \infty$. Then, we define

$$\Psi(x, \gamma) := \varphi(x) + \varphi(\gamma(0)) + \int_0^T \frac{\psi(\dot{\gamma})}{1 + |\gamma|}\, dt$$

and notice that the coercivity of Ψ follows immediately from Lemma 31 below. Then, using assumption (a), we obtain:

$$\begin{aligned}
\int_{\mathbb{R}^d \times \Gamma_T} \Psi(x, \gamma)\, d\boldsymbol{\eta}_h &= \int_{\mathbb{R}^d} \left(2\varphi(x) + \int_0^T \frac{\psi(\dot{\boldsymbol{X}}^h(t, x))}{1 + |\boldsymbol{X}^h(t, x)|}\, dt \right) d\bar{\mu}(x) \\
&= 2 \int_{\mathbb{R}^d} \varphi\, d\bar{\mu} + \int_0^T \int_{\mathbb{R}^d} \frac{\psi(\boldsymbol{b}^h(t, \boldsymbol{X}^h(t, x)))}{1 + |\boldsymbol{X}^h(t, x)|}\, d\bar{\mu}(x) dt \\
&\leq 2 \int_{\mathbb{R}^d} \varphi\, d\bar{\mu} + C \int_0^T \int_{\mathbb{R}^d} \frac{\psi(\boldsymbol{b}^h)}{1 + |y|} dy\, dt.
\end{aligned}$$

Therefore the integrals of Ψ are uniformly bounded.
Step 2. (The limit flow belongs to $\mathscr{L}$) Let now $\boldsymbol{\eta}$ be a narrow limit point of $\boldsymbol{\eta}_h$ along some subsequence that, for notational simplicity, will not be relabelled. Let us show first that $\mu_t^{\boldsymbol{\eta}} = (e_t)_\# \boldsymbol{\eta}$ is representable as $w_t \mathscr{L}^d$ with w_t belonging to $\mathscr{L}$. Indeed, assumption (a) gives

$$\mu_t^{\boldsymbol{\eta}_h} = \boldsymbol{X}^h(t, \cdot)_\# \bar{\mu} = w_t^h \mathscr{L}^d \quad \text{with} \quad \|w_t^h\|_\infty \leq C. \tag{29}$$

Therefore, as $\mu_t^{\boldsymbol{\eta}}$ is the narrow limit of $\mu_t^{\boldsymbol{\eta}_h}$, the same property is preserved in the limit for some $w \in L^\infty$. Moreover, the narrow continuity of $t \mapsto \mu_t^{\eta}$ immediately yields the w^*-continuity of $t \mapsto w_t$ and this proves that $w_t \in \mathscr{L}$.

Step 3. ($\boldsymbol{\eta}$ is concentrated on solutions of the ODE) Next we show that $\boldsymbol{\eta}$ is concentrated on the class of solutions of the ODE. Let $\bar{t} \in [0,T]$, $\chi \in C_c^\infty(\mathbb{R}^d)$ with $0 \le \chi \le 1$, $\boldsymbol{c} \in L^1\left([0,\bar{t}]; L^\infty(\mathbb{R}^d)\right)$, with $\boldsymbol{c}(t,\cdot)$ continuous in $\mathbb{R}^d$ for $\mathscr{L}^1$-a.e. $t \in [0,\bar{t}]$, and define

$$\Phi_c^{\bar{t}}(x,\gamma) := \chi(x) \frac{\left|\gamma(\bar{t}) - x - \int_0^{\bar{t}} \boldsymbol{c}(s,\gamma(s))\, ds\right|}{1 + \sup\limits_{[0,\bar{t}]} |\gamma|^{d+2}}.$$

It is immediate to check that $\Phi_{\boldsymbol{c}}^{\bar{t}} \in C_b(\mathbb{R}^d \times \Gamma_T)$, so that

$$\begin{aligned}
\int_{\mathbb{R}^d \times \Gamma_T} \Phi_{\boldsymbol{c}}^{\bar{t}}\, d\boldsymbol{\eta} &= \lim_{h\to\infty} \int_{\mathbb{R}^d \times \Gamma_T} \Phi_{\boldsymbol{c}}^{\bar{t}}\, d\eta_h \\
&= \lim_{h\to\infty} \int_{\mathbb{R}^d} \chi(x) \frac{\left|\int_0^{\bar{t}} \boldsymbol{b}_h(s, \boldsymbol{X}^h(s,x)) - \boldsymbol{c}(s, \boldsymbol{X}^h(s,x))\, ds\right|}{1 + \sup\limits_{[0,\bar{t}]} |\boldsymbol{X}^h(\cdot,x)|^{d+2}}\, d\bar{\mu}(x) \\
&\le \limsup_{h\to\infty} \int_{\mathbb{R}^d} \int_0^{\bar{t}} \frac{\left|\boldsymbol{b}_h(s, \boldsymbol{X}^h(s,x)) - \boldsymbol{c}(s, \boldsymbol{X}^h(s,x))\right|}{1 + |\boldsymbol{X}^h(s,x)|^{d+2}}\, ds d\bar{\mu}(x) \\
&\le C \limsup_{h\to\infty} \int_0^{\bar{t}} \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_h(s,y) - \boldsymbol{c}(s,y)|}{1 + |y|^{d+2}}\, ds dy \\
&= C \int_0^{\bar{t}} \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}(s,y) - \boldsymbol{c}(s,y)|}{1 + |y|^{d+2}}\, ds dy.
\end{aligned}$$

Now, as (by lower semicontinuity)

$$\frac{|\boldsymbol{b}|}{1 + |x|} \in L^1\left([0,T]; L^1(\mathbb{R}^d)\right) + L^\infty\left([0,T]; L^\infty(\mathbb{R}^d)\right),$$

we can find a sequence of vector fields $\boldsymbol{c}_h(t,x)$ continuous with respect to x and satisfying the assumptions of Lemma 28. Indeed, writing $\boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2 = (\boldsymbol{f}_1 + \boldsymbol{f}_2)(1 + |x|)$, with

$$|\boldsymbol{f}_1| \in L^1\left([0,T]; L^1(\mathbb{R}^d)\right), \qquad |\boldsymbol{f}_2| \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right),$$

we define $\boldsymbol{c} = (\boldsymbol{c}_h)_1 + (\boldsymbol{c}_h)_2$, with

$$(\boldsymbol{c}_h)_i := ((1 + |x|)\boldsymbol{f}_i) * \rho_{\epsilon_h} \qquad i = 1,\, 2$$

with $\epsilon_h \downarrow 0$. It is not hard to show that

$$|(\boldsymbol{c}_h)_i| \le 2(1 + |x|)(|\boldsymbol{f}_i| * \rho_{\epsilon_h}) \qquad i = 1, 2$$

if the support of the convolution kernel ρ is contained in the unit ball. Therefore

$$\frac{|(\boldsymbol{c}_h)_1|}{1+|x|} \quad \text{is bounded and equi-integrable in } L^1\left([0,T]\times\mathbb{R}^d\right), \tag{30}$$

$$\sup_h \left\| \frac{|(\boldsymbol{c}_h)_2(t,\cdot)|}{1+|x|} \right\|_\infty \in L^1(0,T). \tag{31}$$

In the previous estimate we can now choose $\boldsymbol{c} = \boldsymbol{c}_h$ and use Lemma 28 again to obtain

$$\lim_{h\to\infty} \int_{\mathbb{R}^d\times\Gamma_T} \chi(x) \frac{\left|\gamma(\bar{t}) - x - \int_0^{\bar{t}} \boldsymbol{c}_h(s,\gamma(s))\,ds\right|}{1+\sup\limits_{[0,\bar{t}]}|\gamma|^{d+2}}\, d\boldsymbol{\eta} = 0. \tag{32}$$

Now, using the upper bound (29) and Lemma 28 once more we get

$$\begin{aligned}
&\limsup_{h\to\infty} \int_{\mathbb{R}^d\times\Gamma_T} \frac{\int_0^{\bar{t}} |\boldsymbol{c}_h(s,\gamma(s)) - \boldsymbol{b}(s,\gamma(s))|\,ds}{1+\sup\limits_{[0,\bar{t}]}|\gamma|^{d+2}}\, d\boldsymbol{\eta} \qquad (33)\\
\leq\ & \limsup_{h\to\infty} \int_0^T \int_{\mathbb{R}^d\times\Gamma_T} \frac{|\boldsymbol{c}_h(s,\gamma(s)) - \boldsymbol{c}(s,\gamma(s))|}{1+|\gamma(s)|^{d+2}}\, d\boldsymbol{\eta}\, ds \\
\leq\ & C \limsup_{h\to\infty} \int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{c}_h - \boldsymbol{b}|}{1+|x|^{d+2}}\, dxds = 0.
\end{aligned}$$

Hence, from (32) and (33) and Fatou's lemma we infer that for $\chi\boldsymbol{\eta}$-a.e. (x,γ) there is a subsequence $\epsilon_{i(l)}$ such that

$$\lim_{l\to\infty} \left|\gamma(\bar{t}) - x - \int_0^{\bar{t}} \boldsymbol{b}^{\epsilon_{i(l)}}(s,\gamma(s))\,ds\right| + \int_0^{\bar{t}} |\boldsymbol{b}^{\epsilon_{i(l)}}(s,\gamma(s)) - \boldsymbol{b}(s,\gamma(s)|\,ds = 0,$$

so that

$$\gamma(\bar{t}) = x + \int_0^{\bar{t}} \boldsymbol{b}(s,\gamma(s))\,ds.$$

Choosing a sequence of cut-off functions χ_R and letting t vary in $\mathbb{Q}\cap[0,T]$ we obtain that (x,γ) solve the ODE in $[0,T]$ for $\boldsymbol{\eta}$-a.e. (x,γ).

Step 4. (Conclusion) As we are assuming that the uniqueness property holds in $\mathscr{L}$ for the continuity equation relative to $\boldsymbol{b}$, we are now in the position of applying Theorem 18, which says that under these conditions necessarily

$$\boldsymbol{\eta} = (x, \boldsymbol{X}(\cdot,x))_\#\,\bar{\mu}$$

for a suitable map $x \mapsto \boldsymbol{X}(\cdot,x)$. Clearly, by the concentration property of $\boldsymbol{\eta}$, $\boldsymbol{X}(\cdot,x)$ has to be a solution of the ODE for $\bar{\mu}$-a.e. x. This proves that $\boldsymbol{X}$ is the $\mathscr{L}$-Lagrangian flow relative to $\boldsymbol{b}$. The convergence in measure of $\boldsymbol{X}_h$ to $\boldsymbol{X}$ follows by a general principle, stated in Lemma 32 below. $\square$

Lemma 31 (A coercive functional in Γ_T) *Let φ, $\psi : \mathbb{R}^d \to \mathbb{R}$ and let*

$$\Phi(\gamma) := \varphi(\gamma(0)) + \int_0^T \frac{\psi(\dot\gamma)}{1+|\gamma|}\,dt$$

be defined on the subspace of Γ_T made by absolutely continuous maps, and set equal to $+\infty$ outside. If $\varphi(x)$, $\psi(x) \to +\infty$ as $|x| \to \infty$ then all sublevel sets $\{\Phi \le c\}$, $c \in \mathbb{R}$, are compact in Γ_T.

Proof. Let γ_n be such that $\Phi(\gamma_n)$ is bounded and notice that necessarily $|\gamma_n(0)|$ is bounded, by the assumption on φ. By integration of the ODE

$$\frac{d}{dt}\ln(1+|\gamma(t)|) = \frac{\gamma(t)}{|\gamma(t)|} \cdot \frac{\dot\gamma(t)}{1+|\gamma(t)|}$$

one obtains that also $\sup_{[0,T]}|\gamma_n|$ is uniformly bounded. As a consequence the factor $1/(1+|\gamma_n|)$ inside the integral part of Φ can be uniformly estimated from below, and therefore (due to the more than linear growth at infinity of ψ) the sequence $|\dot\gamma_n|$ is equi-integrable in $L^1\left((0,T);\mathbb{R}^d\right)$. As a consequence the sequence (γ_n) is relatively compact in Γ_T. □

Lemma 32 (Narrow convergence and convergence in measure) *Let v_h, $v : X \to Y$ be Borel maps and let $\bar\mu \in \mathscr{M}_+(X)$. Then $v_h \to v$ in $\bar\mu$-measure iff*

$$(x, v_h(x))_\#\bar\mu \text{ converges to } (x, v(x))_\#\bar\mu \text{ narrowly in } \mathscr{M}_+(X \times Y).$$

Proof. If $v_h \to v$ in $\bar\mu$-measure then $\varphi(x, v_h(x))$ converges in $L^1(\bar\mu)$ to $\varphi(x, v(x))$, and we immediately obtain the convergence of the push-forward measures. Conversely, let $\delta > 0$ and, for any $\epsilon > 0$, let $w \in C_b(X;Y)$ be such that $\bar\mu(\{v \ne w\}) \le \epsilon$. We define

$$\varphi(x,y) := 1 \wedge \frac{d_Y(y, w(x))}{\delta} \in C_b(X \times Y)$$

and notice that

$$\bar\mu\left(\{v \ne w\}\right) + \int_{X\times Y} \varphi\, d(x, v_h(x))_\#\bar\mu \ge \bar\mu(\{d_Y(v, v_h) > \delta\}),$$

$$\int_{X\times Y} \varphi\, d(x, v(x))_\#\bar\mu \le \bar\mu(\{w \ne v\}).$$

Taking into account the narrow convergence of the push-forward we obtain that

$$\limsup_{h\to\infty} \bar\mu(\{d_Y(v, v_h) > \delta\}) \le 2\bar\mu(\{w \ne v\}) \le 2\epsilon\,;$$

since ϵ is arbitrary the proof is achieved. □

The renormalization Theorem 24 has been extended in [6] to the case of a BV dependence w.r.t. the spatial variables, but still assuming that

$$D \cdot \boldsymbol{b}_t \ll \mathscr{L}^d \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{34}$$

Theorem 33 *Let $\boldsymbol{b} \in L^1_{\rm loc}\left((0,T); BV_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ be satisfying (34). Then any distributional solution $w \in L^\infty_{\rm loc}\left((0,T)\times\mathbb{R}^d\right)$ of*

$$\frac{d}{dt}w + D_x \cdot (\boldsymbol{b}w) = c \in L^1_{\rm loc}\left((0,T)\times\mathbb{R}^d\right)$$

is a renormalized solution.

A self contained proof of this result, slightly simpler than the one given in the original paper [6], is given in [7] and [8]. The original argument in [6] was indeed based on deep result of G. Alberti [2], saying that for a $BV_{\rm loc}$ function $u : \mathbb{R}^d \to \mathbb{R}^m$ the matrix $M(x)$ in the polar decomposition $Du = M|Du|$ has rank 1 for $|D^s u|$-a.e. x, i.e. there exist unit vectors $\xi(x) \in \mathbb{R}^d$ and $\eta(x) \in \mathbb{R}^m$ such that $M(x)z = \eta(x)\langle z, \xi(x)\rangle$. However, we observe that in the application of this result to the Keyfitz–Kranzer system [11], [9] the vector field $\boldsymbol{b}$ is of the form $\boldsymbol{f}(\rho)$ with ρ scalar and $\boldsymbol{f} \in C^1$ vectorial, so the rank-one structure of the distributional derivative (as a whole) is easy to check. Analogously, in the case of the semi-geostrophic system considered in [53], the vector field is a monotone map, and for this class of BV functions a much simpler proof of the rank-one property is available [3].
As in the Sobolev case we can now obtain from the general theory given in Section 3 *existence and uniqueness* of $\mathscr{L}$-Lagrangian flows, with $\mathscr{L} = L^\infty(L^1) \cap L^\infty(L^\infty)$: we just replace in the statement of Theorem 26 the assumption $\boldsymbol{b} \in L^1\left([0,T]; W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ with $\boldsymbol{b} \in L^1\left([0,T]; BV_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$, assuming as usual that $D\cdot\boldsymbol{b}_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in [0,T]$.
Analogously, by applying Theorem 29 we obtain *stability* of $\mathscr{L}$-Lagrangian flows when $\boldsymbol{b}$ is as in Theorem 33.

6 Measure-theoretic differentials

In this section we introduce some weak differentiability notions, based on measure-theoretic limits of difference quotients, and we compare them (all results of this section are taken from [17]). An important remark is that none of these concepts gives additional informations on the derivative in the sense of distributions. Conversely, whenever the derivative in the sense of distributions is a measure with locally finite variation, the map is $\mathscr{L}^d$-a.e. approximately differentiable.
We recall that a sequence of measurable maps (f_k) defined on an open set $\Omega \subset \mathbb{R}^d$ is said to converge locally in measure to a measurable map f if

$$\lim_{k\to\infty} \mathscr{L}^d\left(\{x \in K : \ |f_k(x) - f(x)| > \epsilon\}\right) = 0$$

for any compact set $K \subset A$ and any $\epsilon > 0$.
The following simple lemma will be used in many occasions.

Lemma 34 *Let f_k, $f:\ \Omega\times\mathbb{R}^d\to\mathbb{R}$ be measurable functions. Suppose that, for each $y\in\mathbb{R}^d$, $f_k(\cdot,y)\to f(\cdot,y)$ locally in measure in Ω as $k\to\infty$. Then $f_k\to f$ locally in measure in $\Omega\times\mathbb{R}^d$ as $k\to\infty$.*

Proof. Let $K\subset\Omega$ be a a compact set. We fix $\gamma>0$ and set

$$g_k(y)=\mathscr{L}^d\left(\{x\in K:\ |f_k(x,y)-f(x,y)|\geq\gamma\}\right).$$

Then from the dominated convergence theorem we infer that

$$g_k\to 0\quad\text{ in } L^1_{\rm loc}(\mathbb{R}^d).$$

In particular, Fubini's theorem gives that $\mathscr{L}^{2d}\big((K\times B)\cap\{|f_k-f|>\gamma\}\big)\to 0$ for any ball $B\subset\mathbb{R}^d$. □

As a consequence of the Lusin theorem, it is easy to check that for any measurable function $f:\Omega\to\mathbb{R}$ the functions $f(x+h)$ converge to f locally in measure as $h\to 0$. Therefore it is natural to study the behaviour, still with respect to local convergence in measure, of the difference quotients. This leads to the following definition.

Definition 35 (Fréchet and Gâteaux derivative in measure) *Let $f:\Omega\subset\mathbb{R}^d\to\mathbb{R}$ be a measurable function. We say that $g:\Omega\to\mathbb{R}^d$ is the (Fréchet) derivative in measure of f if*

$$\lim_{h\to 0}\frac{f(x+h)-f(x)-g(x)\cdot h}{|h|}=0\quad\textit{locally in measure in }\Omega.$$

Analogously, g is called Gâteaux derivative in measure if the difference quotients above tend to 0 locally in measure in Ω along all lines passing through the origin.

Another differentiability condition, that we call directional differentiability in measure, involves an averaging procedure also on the direction. It appeared first in [73] in connection with the differentiability properties of the flow associated to Sobolev vector fields (see the next section) and it can be stated as follows.

Definition 36 (Directional differentiability in measure) *We say that $f:\Omega\to\mathbb{R}$ is directionally differentiable in measure if there exists $W:\Omega\times\mathbb{R}^d\to\mathbb{R}$ such that*

$$\frac{f(x+ry)-f(x)-rW(x,y)}{r}\to 0\qquad\textit{locally in measure in }\Omega\times\mathbb{R}^d\textit{ as }h\to 0.$$

The following result surprisingly shows that these three concepts are equivalent.

Theorem 37 *Let $f:\Omega\subset\mathbb{R}^d\to\mathbb{R}$ be a measurable function. Then the following assertions are equivalent:*

(i) f is (Fréchet) differentiable in measure in Ω;

(ii) f is Gâteaux differentiable in measure in Ω;

(iii) f is directionally differentiable in measure in Ω.

Moreover, the derivative in measure g is linked to the directional derivative in measure W by

$$g(x) \cdot y = W(x,y) \qquad \textit{for } \mathscr{L}^{2d}\textit{-a.e. } (x,y) \in \Omega \times \mathbb{R}^d.$$

Proof. We start with some preliminary remarks. It is obvious that if g is the derivative in measure of f then g is also the Gâteaux derivative in measure of f, and that these derivatives in measure are unique, up to $\mathscr{L}^d$-negligible sets.
Moreover, Gâteaux differentiability in measure implies directional differentiability in measure, with $W(x,y) = g(x)y$: this fact is an immediate consequence of Lemma 34.
The harder implication is the one from directional differentiability in measure to Fréchet differentiability in measure. We need the following lemma (see [71], [56]): *let $f : \mathbb{R}^d \to \mathbb{R}$ be a measurable function such that*

$$f(y+z) = f(y) + f(z) \text{ for } \mathscr{L}^{2d}\text{-a.e. } (y,z) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Then there exists $a \in \mathbb{R}^d$ such that $f(y) = a \cdot y$ for $\mathscr{L}^d$-a.e. $y \in \mathbb{R}^d$.
Step 1. Using the above mentioned lemma, we show the "a.e. linearity" of $W(x,\cdot)$. We define

$$\begin{aligned}
F_\delta(x,y,z) &= \frac{f(x+\delta y) - f(x) - \delta W(x,y)}{\delta}, \\
G_\delta(x,y,z) &= \frac{f(x+\delta y+\delta z) - f(x+\delta y) - \delta W(x+\delta y, z)}{\delta}, \\
H_\delta(x,y,z) &= \frac{-f(x+\delta y+\delta z) + f(x) + \delta W(x, y+z)}{\delta}.
\end{aligned}$$

Then

$$F_\delta \to 0 \quad \text{as } \delta \to 0, \tag{35}$$

$$G_\delta \to 0 \quad \text{as } \delta \to 0, \tag{36}$$

$$H_\delta \to 0 \quad \text{as } \delta \to 0 \tag{37}$$

locally in measure with respect to (x,y,z). Indeed, (35) and (37) are easy applications of Lemma 34. To verify (36) we need also the following shift argument. Let $\Omega'' \subset\subset \Omega' \subset\subset \Omega$. For small δ we obtain

$$\begin{aligned}
&\mathscr{L}^{3d}\Big(\big\{(x,y,z) \in \Omega'' : \ G_\delta(x,y,z)) \geq \gamma\big\}\Big) \\
&\quad = \mathscr{L}^{3d}\Big(\big\{(x,y,z) \in \Omega'' : \ |f(x+\delta y+\delta z) - f(x+\delta y) - \delta W(x+\delta y, z)| \geq \gamma\delta\big\}\Big) \\
&\quad \leq \mathscr{L}^{3d}\Big(\big\{(x',y,z) \in \Omega' : |f(x'+\delta z) - f(x') - \delta W(x',z)| \geq \gamma\delta\big\}\Big) \to 0
\end{aligned}$$

where we used the change of variables $x' = x + \delta y$. This proves (36). From Lemma 34 we infer that

$$W(x+\delta y, z) - W(x,z) \to 0 \quad \text{as } \delta \to 0.$$

Adding the three identities we obtain that

$$W(x,y+z)-W(x,y)-W(x,z)=0 \quad \text{for } \mathscr{L}^{3d}\text{-a.e. } (x,y,z)\in\Omega\times\mathbb{R}^d\times\mathbb{R}^d.$$

Therefore, for $\mathscr{L}^d$-a.e. x, $W(x,\cdot)$ is representable as a linear function. This concludes Step 1.
Step 2. Let $\boldsymbol{\alpha}_d$ be the measure of the unit ball in $\mathbb{R}^d$. We fix a set $\Omega'\subset\subset\Omega$ and ϵ, $\gamma>0$. We find $\delta_1>0$ such that

$$B(x,2\delta_1)\subset\Omega, \qquad x\in\Omega'.$$

Consider the sets

$$\begin{aligned} A_\delta(x) = \{y\in B(0,2)\colon |f(x+\delta y)-f(x)-\delta g(x)\cdot y|\geq\gamma\delta\},\\ 0<\delta<\delta_1, \quad x\in\Omega'. \end{aligned}$$

Since $(x,y)\mapsto g(x)\cdot y$ is a directional derivative of f in measure, by the Fubini theorem,

$$\int_{\Omega'}\mathscr{L}^d\left(A_\delta(x)\right)\,dx\to 0 \quad \text{as } \delta\to 0.$$

Therefore there exists $\delta_2>0$ such that for $0<\delta<\delta_2$ we have

$$\mathscr{L}^d\left(\left\{x\in\Omega'\colon \mathscr{L}^d(A_\delta(x))\geq\frac{\boldsymbol{\alpha}_d}{4}\right\}\right)\leq\epsilon.$$

By Lusin theorem, there exists a compact set $K\subset\Omega$ such that $\mathscr{L}^d(\Omega\setminus K)<\epsilon$ and $g|_K$ is continuous. By the uniform continuity, there exists $\delta_3>0$ such that for $0<\delta<\delta_3$ we have

$$x,\,x'\in K,\ |x-x'|<\delta \quad\Longrightarrow\quad |g(x)-g(x')|<\gamma.$$

Let $h\in\mathbb{R}^d$, $|h|=r<\frac{1}{2}\min\{\delta_1,\delta_2,\delta_3\}$. Consider the sets

$$\begin{aligned} E &:= \{x\in\Omega':\ \mathscr{L}^d(A_r(x))\geq\tfrac{1}{4}\boldsymbol{\alpha}_d\},\\ E' &:= \{x\in\Omega':\ \mathscr{L}^d(A_r(x+h))\geq\tfrac{1}{4}\boldsymbol{\alpha}_d\},\\ F &:= \Omega'\setminus K,\\ F' &:= \{x\in\Omega':\ x+h\notin K\}. \end{aligned}$$

Then

$$|E\cup E'\cup F\cup F'|\leq 4\epsilon.$$

Let $x\in\Omega\setminus(E\cup E'\cup F\cup F')$. Then both x and $x+h$ belong to K and

$$\begin{aligned} \mathscr{L}^d\left(B(0,1)\cap A_r(x)\right)\leq\tfrac{1}{4}\boldsymbol{\alpha}_d,\\ \mathscr{L}^d\left(\{y\in B(0,1):\ y-\tfrac{h}{r}\in A_r(x+h)\}\right)\leq\tfrac{1}{4}\boldsymbol{\alpha}_d. \end{aligned}$$

Thus there exists $y\in B(0,1)\setminus A_r(x)$ such that $y':=y-\frac{h}{r}\notin A_r(x+h)$. Since $|y'|<2$, the fact that $y'\notin A_r(x+h)$ means that $x+h+ry'\in\Omega$ and

$$|f(x+h+ry')-f(x+h)-rg(x+h)\cdot y'|\leq\gamma r.$$

Since $h + ry' = ry$, we have

$$\begin{aligned} & |f(x+h)-f(x)-g(x)\cdot h| \\ \leq\ & |f(x+h)-f(x+h+ry')+rg(x+h)\cdot y'| \\ & +|f(x+ry)-f(x)-rg(x)\cdot y| + |r(g(x)-g(x+h))\cdot y'| \\ \leq\ & \gamma r + \gamma r + \gamma r|y'| \leq 4|h|\gamma\,. \end{aligned}$$

This shows that

$$\frac{f(x+h)-f(x)-g(x)\cdot h}{|h|} \to 0$$

in measure w.r.t. x as $h \to 0$. $\square$

Now we introduce another more classical weak differentiability property (extensively studied, for instance, in [68]). It has still a measure-theoretic character, but unlike differentiability in measure it has a pointwise meaning.

Definition 38 (Approximate differentiability) *Let $f : \Omega \to \mathbb{R}$ be a measurable function and let $x \in \Omega$. We say that $a \in \mathbb{R}^d$ is an approximate derivative of f at x if*

$$\left\{h\colon \frac{|f(x+h)-f(x)-a\cdot h|}{|h|} > \epsilon\right\} \quad \text{has zero Lebesgue density at } 0 \text{ for any } \epsilon > 0.$$

As we are concerned here with convergence in measure, it is worth mentioning that approximate differentiability at x is equivalent to the convergence

$$\lim_{\delta\to 0} \frac{f(x+\delta y)-f(x)-\delta a\cdot y}{\delta} = 0 \quad \text{locally in measure w.r.t. } y \in \mathbb{R}^d.$$

The following proposition shows that functions which are approximately differentiable on a measurable set A essentially coincide with functions that can be approximated, in the Lusin sense, by Lipschitz maps.

Theorem 39 (Lusin theorem for approximately differentiable maps) *Let $f : \Omega \to \mathbb{R}$. Assume that there exists a sequence of measurable sets $A_n \subset \Omega$ such that $\mathscr{L}^d(\Omega \setminus \cup_n A_n) = 0$ and $f|_{A_n}$ is Lipschitz for any n. Then f is approximately differentiable at $\mathscr{L}^d$-a.e. $x \in \Omega$.*
Conversely, if f is approximately differentiable at all points of $\Omega' \subset \Omega$, we can write Ω' as a countable union of sets A_n such that $f|_{A_n}$ is Lipschitz for any n (up to a redefinition in a $\mathscr{L}^d$-negligible set).

Proof. With no loss of generality we can assume that the sets A_n are pairwise disjoint. By Mc Shane Lipschitz extension theorem we can find Lipschitz functions $g_n : \mathbb{R}^d \to \mathbb{R}$ extending $f|_{A_n}$. By Lebesgue differentiation theorem and Rademacher theorem, $\mathscr{L}^d$-a.e. $x \in A_n$ is both a point of density 1 of A_n and a differentiability point of g_n. We claim that at any of these points x the function f is approximately differentiable, with approximate differentiable equal to $\nabla g_n(x)$. Indeed, it suffices to notice that the difference quotients of f and of g_n may differ only on $\Omega \setminus A_n$, that has zero density at x.
Clearly, as $\mathscr{L}^d$-a.e. $x \in \Omega$ has this property for some n, so this proves the $\mathscr{L}^d$-a.e. approximate differentiability of f.
The converse statement is proved in Theorem 3.1.16 of [68]. $\square$

The following theorem shows that, among all the differentiability properties considered in this section, $\mathscr{L}^d$-a.e. approximate differentiability is the stronger one. Even on the real line, an example built in [17] shows that differentiability in measure does not imply $\mathscr{L}^1$-a.e. approximate differentiability (in fact, the function built in [17] is nowhere approximately differentiable).

Theorem 40 *Suppose that $f : \Omega \to \mathbb{R}$ is approximately differentiable $\mathscr{L}^d$-a.e. in Ω. Then f is differentiable in measure.*

Proof. Let g be the approximate derivative of f. By the previous result there exist Lipschitz functions f_j and pairwise disjoint measurable sets $A_j \subset \Omega$ such that

$$
\begin{aligned}
f_j &= f \text{ on } A_j, \\
\nabla f_j &= g \text{ on } A_j, \\
\mathscr{L}^d\Big(\Omega \setminus \bigcup_j A_j\Big) &= 0.
\end{aligned}
$$

We fix $\Omega' \subset\subset \Omega$ and $\gamma > 0$ and find $\delta > 0$ such that $B(x, \delta) \subset \Omega$ for each $x \in \Omega'$. We set

$$E(h) := \{x \in \Omega' \colon |f(x+h) - f(x) - g(x)h| \geq \gamma |h|\}, \quad h \in B(0, \delta).$$

Then

$$E(h) \subset N \cup \bigcup_j (E_j(h) \cup F_j(h)),$$

where $\mathscr{L}^d(N) = 0$ and

$$
\begin{aligned}
E_j(h) &= \big\{x \in \Omega' \cap A_j : \ x + h \notin A_j\big\}, \\
F_j(h) &= \big\{x \in \Omega' \cap A_j : \ |f_j(x+h) - f_j(x) - \nabla f_j(x)h| \geq \gamma |h|\big\}.
\end{aligned}
$$

Then

$$\mathscr{L}^d(E_j(h)) \to 0 \quad \text{as } h \to 0$$

by the L^1_{loc} continuity of translations. Using the differentiability of f_j on A_j, we obtain that also

$$\mathscr{L}^d(F_j(h)) \to 0 \quad \text{as } h \to 0.$$

Choose $\epsilon > 0$. Since $E_j(h) \cup F_j(h) \subset A_j$ and $\sum_j \mathscr{L}^d(A_j) \leq \mathscr{L}^d(\Omega) < \infty$, we can find an index m independent of h such that

$$\sum_{j=m+1}^{\infty} \mathscr{L}^d\left(E_j(h) \cup F_j(h)\right) \leq \epsilon.$$

Then

$$\limsup_{h \to 0} \mathscr{L}^d(E(h)) \leq \sum_{j=1}^{m} \limsup_{h \to 0} \mathscr{L}^d\left(E_j(h) \cup F_j(h)\right) + \epsilon = \epsilon.$$

It follows that $\mathscr{L}^d(E(h)) \to 0$ as required. $\square$

7 Differentiability of the flow in the $W^{1,1}$ case

In this section we discuss the differentiability properties of the $\mathscr{L}$-Lagrangian flow associated to $W^{1,1}$ vector fields, briefly describing the results obtained in [73]. Notice that no differentiability property is presently known in the BV case; on the other hand, in the $W^{1,p}$ case, with $p > 1$, much stronger results are available [48], and we will present them in the next section.
The main theorem in [73] is the following:

Theorem 41 *Let $\boldsymbol{b} \in L^1\left([0,T]; W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ be satisfying*

(i) $\dfrac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T]; L^1(\mathbb{R}^d)\right) + L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right)$;

(ii) $[\operatorname{div} \boldsymbol{b}_t] \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right)$;

and let $\boldsymbol{X}(t,x)$ be the corresponding $\mathscr{L}$-Lagrangian flow, given by Theorem 26. Then for all $t \in [0,T]$ there exists a measurable function $\boldsymbol{W}_t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\frac{\boldsymbol{X}(t,x+\epsilon y) - \boldsymbol{X}(t,x) - \epsilon \boldsymbol{W}_t(x,y)}{\epsilon} \to 0 \quad \textit{locally in measure in } \mathbb{R}^d_x \times \mathbb{R}^d_y$$

as $\epsilon \downarrow 0$.

The result actually stated in [73] is slightly stronger, as the convergence above is also shown to be uniform with respect to time. Having in mind the terminology and the results of the previous section, the result can also be rephrased as follows: for all $t \in [0,T]$ there exists a matrix-valued measurable function $\boldsymbol{G}_t : \mathbb{R}^d \to M^{d\times d}$ (i.e. the derivative in measure) such that

$$\frac{\boldsymbol{X}(t,x+h) - \boldsymbol{X}(t,x) - \boldsymbol{G}_t(x)h}{|h|} \to 0 \qquad \text{locally in measure in } \mathbb{R}^d_x \text{ as } h \to 0.$$

The link between $\boldsymbol{G}_t$ and $\boldsymbol{W}_t$ is given by $\boldsymbol{G}_t(x)(y) = \boldsymbol{W}_t(x,y)$ for $\mathscr{L}^{2d}$-a.e. (x,y). So, $\boldsymbol{G}_t$ can be interpreted as the derivative of the flow map $\boldsymbol{X}(t,\cdot)$.
The strategy of the proof in [73] is to look at the behaviour of the $2d$-dimensional flows $\boldsymbol{Y}^\epsilon$ arising from the difference quotients of $\boldsymbol{X}$:

$$\boldsymbol{Y}^\epsilon(t,x,y) := \frac{\boldsymbol{X}(t,x+\epsilon y) - \boldsymbol{X}(t,x)}{\epsilon}.$$

It is immediate to check (see also [15]) that $\boldsymbol{Y}^\epsilon$ are $\mathscr{L}$-Lagrangian flows relative to the vector fields $\boldsymbol{B}^\epsilon$ defined by

$$\boldsymbol{B}^\epsilon(t,x,y) = \left(\boldsymbol{b}(t,x), \frac{\boldsymbol{b}(t,x+\epsilon y) - \boldsymbol{b}(t,x)}{\epsilon}\right).$$

Therefore, it is natural to expect that their limit $\boldsymbol{Y}$ (if any) should be a flow relative to the limit vector field

$$\boldsymbol{B}(t,x,y) = \left(\boldsymbol{b}(t,x), \nabla_x \boldsymbol{b}(t,x) y\right).$$

Notice that $\operatorname{div} \boldsymbol{B}_t(x,y) = 2\operatorname{div} \boldsymbol{b}_t(x)$, and therefore

$$\operatorname{div} \boldsymbol{B}_t \in L^1\left([0,T]; L^\infty(\mathbb{R}^{2d})\right).$$

On the other hand, the last d components of $\boldsymbol{B}$ have no regularity with respect to the x-variable. However, in this special case, an anisotropic smoothing argument (by a regularization in the x variable much faster than the one in the y variable), based on the special structure of this vector field (see also [25], [74]), still guarantees that bounded solutions to the continuity equation with velocity field $\boldsymbol{B}$ are renormalizable. Moreover, using the renormalization property, a variant of the argument used in Theorem 23 shows that the continuity equation with $\boldsymbol{B}$ as vector field satisfies the comparison principle in the class

$$\mathscr{L}^* := \mathscr{L} \cap L^\infty\left([0,T]; L^\infty_{\mathrm{loc}}(\mathbb{R}^d_x; L^1(\mathbb{R}^d_y))\right),$$

where $\mathscr{L}$ is defined as in (18) (in $2d$ space dimensions). The reason for this restriction to the smaller space $\mathscr{L}^*$ is the fact that $|\boldsymbol{B}|/(1+|x|+|y|)$ in general does not belong to

$$L^1\left([0,T]; L^1(\mathbb{R}^d_x \times \mathbb{R}^d_y)\right) + L^1\left([0,T]; L^\infty(\mathbb{R}^d_x \times \mathbb{R}^d_y)\right),$$

because the last d components do not tend to 0 as $|y| \to \infty$ while x is kept fixed (and their limit is possibly unbounded as a function of x). If $\boldsymbol{b}$ satisfies condition (ii) above, the last d components $\boldsymbol{B}^2$ of $\boldsymbol{B}$ satisfy instead

$$\frac{|\boldsymbol{B}^2|}{1+|y|} \in L^1\left([0,T]; L^1_{\mathrm{loc}}(\mathbb{R}^d_x, L^1(\mathbb{R}^d_y) + L^\infty(\mathbb{R}^d_y))\right).$$

For this reason, a weaker growth condition on $\boldsymbol{B}$ turns into a stronger growth condition on w. Then, the renormalization property ensures the well posedness of the continuity equation and therefore (much like as in Theorem 16 and Theorem 23) that the $\mathscr{L}^*$-Lagrangian flow relative to $\boldsymbol{B}$ is unique. A smoothing argument (see [73] for details) proves its existence even within the smaller class $\mathscr{L}^*$. So, denoting by $\boldsymbol{Y}$ the $\mathscr{L}^*$-Lagrangian flow relative to $\boldsymbol{B}$, we can represent it as

$$\boldsymbol{Y}(t,x,y) = (\boldsymbol{X}(t,x), \boldsymbol{W}(t,x,y))$$

for some map $\boldsymbol{W}$. Finally, the same argument used for the existence of $\boldsymbol{Y}$ shows that $\boldsymbol{Y}$ is the limit, with respect to local convergence in measure (uniform w.r.t. time) of $\boldsymbol{Y}^\epsilon$: this is due to the fact that $\boldsymbol{B}^\epsilon$ have properties analogous to the ones of $\boldsymbol{B}$ (in particular they have uniformly bounded divergences) and converge to $\boldsymbol{B}$.
This leads to the proof of Theorem 41.

8 Differentiability and compactness of the flow in the $W^{1,p}$ case

In this section we present some recent results, obtained in [48], relative to the approximate differentiability and to the Lipschitz properties of regular Lagrangian flows associated to $W^{1,p}$

vector fields, with $p > 1$. The first results relative to the approximative differentiability of the flow have been obtained in [15]. An important fact that we want to remark from the beginning is that the approach of [15] and [48] is completely different from the one of [73] described in the previous section: these two papers are not based on the theory of renormalized solutions, but new kind of estimates are introduced. The general idea, which we are going to explain with all the details in the following, is trying to find estimates for the spatial gradient of the flow in terms of bounds on the derivative of the vector field.

We start by recalling some basic facts about the the theory of *maximal functions*. These tools will be used throughout all this section. We begin with the definition of the local maximal function.

Definition 42 (Maximal function) *Let μ be a (vector-valued) measure with locally finite total variation. For every $\lambda > 0$, we define the* (λ-local) maximal function *of μ as*

$$M_\lambda \mu(x) \;=\; \sup_{0<r<\lambda} \frac{|\mu|(B_r(x))}{\mathscr{L}^d(B_r(x))} \qquad\qquad x \in \mathbb{R}^d .$$

When $\mu = f\mathscr{L}^d$, where f is a function in $L^1_{\rm loc}(\mathbb{R}^d;\mathbb{R}^m)$, we will use the notation $M_\lambda f$ for $M_\lambda \mu$.

In the following we will use a lot of times the following two lemmas about maximal functions. Their proof is classical and can be found for example in [87]. The first lemma shows that it is possible to control the L^p norm of a maximal function with the L^p norm of the function itself, in the case $p > 1$; however, this estimate is *false* in the case $p = 1$. The second lemma will be used to estimate the difference quotiens of the vector field by means of the maximal function of the spatial derivative of the vector field.

Lemma 43 *The local maximal function of μ is finite for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$ and, for every $R > 0$, we have*

$$\int_{B_R(0)} M_\lambda f(y)\, dy \le c_{d,R} + c_d \int_{B_{R+\lambda}(0)} |f(y)| \log(2 + |f(y)|)\, dy \quad \forall \lambda > 0 .$$

For $p > 1$ we have

$$\int_{B_R(0)} (M_\lambda f(y))^p\, dy \le c_{d,p} \int_{B_{R+\lambda}(0)} |f(y)|^p\, dy \qquad \forall \lambda > 0 .$$

Lemma 44 *If $u \in BV(\mathbb{R}^d;\mathbb{R}^m)$ then there exists an $\mathscr{L}^d$-negligible set $N \subset \mathbb{R}^d$ such that*

$$|u(x) - u(y)| \le c_d |x-y| \left(M_\lambda Du(x) + M_\lambda Du(y) \right)$$

for $x, y \in \mathbb{R}^d \setminus N$ with $|x-y| \le \lambda$.

We also recall Chebyshev inequality for a measurable function $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$:

$$\mathscr{L}^d\big(\{|f| > t\}\big) \leq \frac{1}{t} \int_{\{|f|>t\}} |f(x)|\, dx \leq \frac{\mathscr{L}^d\big(\{|f| > t\}\big)^{1/q}}{t} \|f\|_{L^p(\Omega)} ,$$

which immediately implies

$$\mathscr{L}^d\big(\{|f| > t\}\big)^{1/p} \leq \frac{\|f\|_{L^p(\Omega)}}{t} . \tag{38}$$

In all this section we will make the following assumptions on the vector field $\boldsymbol{b} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$:

(A) $\boldsymbol{b} \in L^1([0,T]; W^{1,p}(\mathbb{R}^d; \mathbb{R}^d))$ for some $p > 1$;

(B) $\boldsymbol{b} \in L^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$;

(C) $[\mathrm{div}_x \boldsymbol{b}]^- \in L^1([0,T]; L^\infty(\mathbb{R}^d))$.

We remark that condition (B) could be relaxed to

$$\frac{|\boldsymbol{b}(t,x)|}{1+|x|} \in L^1([0,T]; L^1(\mathbb{R}^d)) + L^1([0,T]; L^\infty(\mathbb{R}^d)) ,$$

getting some slightly weaker results (due to the possible unboundedness of the velocity field), but in this introductory presentation we prefer to avoid these technicalities and focus on the case of a uniformly bounded vector field. We also notice that, under this assumption, the global $W^{1,p}$ hypothesis is not essential: thanks to the finite speed of propagation, we could truncate our vector field out of a ball, then getting the same results with just a $W^{1,p}_{\mathrm{loc}}$ hypothesis. However, we prefer to assume this global condition, mainly in order to simplify typographically some estimates. We refer to [48] for these more general hypotheses, the main modification being an estimate of the superlevels of the flow.
We recall that in this context the results of Section 5 read as follow. For every vector field $\boldsymbol{b}$ satisfying assumptions (A), (B) and (C) there exist a unique regular Lagrangian flow $\boldsymbol{X}$, that is a measurable map $\boldsymbol{X} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ which satisfies the following two properties:

(i) for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$ the map $t \mapsto \boldsymbol{X}(t,x)$ is an absolutely continuous solution of $\dot{\boldsymbol{\gamma}}(t) = \boldsymbol{b}(t, \boldsymbol{\gamma}(t))$ with $\boldsymbol{\gamma}(0) = x$ (the ODE if fulfilled in the integral sense or, equivalently, in the a.e. sense);

(ii) there exists a constant C independent of t such that $\boldsymbol{X}(t,\cdot)_\# \mathscr{L}^d \leq C \mathscr{L}^d$ for every $t \in [0,T]$.

For every vector field satisfying assumption (C), the quantity

$$C = \exp\left(\int_0^T \|[\mathrm{div}_x \boldsymbol{b}(t,\cdot)]^-\|_{L^\infty(\mathbb{R}^d)}\, dt \right)$$

satisfies the inequality in the second property of the regular Lagrangian flow. However, we remark the fact that all our estimates will depend on the *compressibility constant* L of $\boldsymbol{X}$, *i.e.* the best constant C for which (ii) holds, rather than on the $L^1(L^\infty)$ norm of $[\mathrm{div}_x \boldsymbol{b}]^-$. This is an important remark in the context of the compactness theorem (see Theorem 48): some compactness results under bounds on the divergence were already available in [61], while one of the merits of this new approach is this weaker requirement.

The starting point of the estimates given in [48] is already present in [15] (and, at least in a formal way, in [61]): in a smooth context, we can control the time derivative $\frac{d}{dt}\log(|\nabla \boldsymbol{X}|)$ with $|\nabla \boldsymbol{b}|(\boldsymbol{X})$. The strategy of [15] allows to make this remark rigorous: it is possible to consider some integral quantities which contain a discretization of the space gradient of the flow and to prove some estimates along the flow. Then, the application of Egorov theorem allows the passage from integral estimates to pointwise estimates on big sets, and from this it is possible to recover Lipschitz regularity on big sets. However, the application of Egorov theorem implies a loss of quantitative informations: this strategy does not allow a control of the Lipschitz constant in terms of the size of the "neglected" set.
Starting from this results, the main point of [48] is a modification of the estimates in such a way that quantitative informations are not lost. We define (for $p > 1$ and $R > 0$) the following integral quantity:

$$A_p(R, \boldsymbol{X}) := \left[\int_{B_R(0)} \left(\sup_{0 \le t \le T} \sup_{0 < r < 2R} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1 \right) dy \right)^p dx \right]^{1/p} . \quad (39)$$

Heuristically, one can view $|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|/r$ as a "discrete gradient" of the flow $\boldsymbol{X}$. Then the quantity $A_p(R, \boldsymbol{X})$ is constructed averaging this discrete gradient over balls of radius r, asking some uniformity with respect to t and r and finally integrating on a bounded set with respect to the second variable. We are now going to give some quantitative estimates of the quantity $A_p(R, \boldsymbol{X})$ in the case when the map $\boldsymbol{X}$ is the regular Lagrangian flow associated to a vector field $\boldsymbol{b}$ satisfying assumptions (A), (B) and (C). The estimate will depend only on the $L^1(L^p)$ norm of the derivative of $\boldsymbol{b}$ and on the compressibility constant L. In all the following computations we will denote by $c_{q_1,\dots,q_n}$ universal constants which depends only on the parameters $q_1, \dots, q_n$ and which can change from line to line. To simplify the notation we will also denote by L^p and $L^1(L^p)$ the global spaces $L^p(\mathbb{R}^d)$ and $L^1([0,T]; L^p(\mathbb{R}^d))$ respectively. Out of the smooth setting, all the computations are easily justified by condition (i) in the definition of regular Lagrangian flow; however, the reader could check the estimates in the smooth case, and then obtain the general case simply by an approximation procedure (based on Theorem 29).

Proposition 45 *Let $\boldsymbol{b}$ be a vector field satisfying assumptions (A), (B) and (C). Denote by $\boldsymbol{X}$ its regular Lagrangian flow and let L be the compressibility constant of the flow. Then we have*

$$A_p(R, \boldsymbol{X}) \le C\left(R, L, \|D_x \boldsymbol{b}\|_{L^1(L^p)}\right) .$$

Proof. For $0 \le t \le T$, $0 < r < 2R$ and $x \in B_R(0)$ define

$$Q(t, x, r) := \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1 \right) dy .$$

With some easy computations we get

$$
\begin{aligned}
\frac{dQ}{dt}(t,x,r) &\leq \fint_{B_r(x)} \left| \frac{d\boldsymbol{X}}{dt}(t,x) - \frac{d\boldsymbol{X}}{dt}(t,y) \right| \left(|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)| + r \right)^{-1} \, dy \\
&= \fint_{B_r(x)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x)) - \boldsymbol{b}(t,\boldsymbol{X}(t,y))|}{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)| + r} \, dy \,. \qquad (40)
\end{aligned}
$$

We now set $\tilde{R} = 4R + 2T\|\boldsymbol{b}\|_\infty$. Since we clearly have $|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)| \leq \tilde{R}$, applying Lemma 44 we can estimate

$$
\begin{aligned}
\frac{dQ}{dt}(t,x,r) &\leq c_d \fint_{B_r(x)} \left(M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) + M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,y)) \right) \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)| + r} \, dy \\
&\leq c_d M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) + c_d \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,y)) \, dy \,.
\end{aligned}
$$

Integrating with respect to the time, we estimate

$$
\sup_{0\leq t\leq T} Q(t,x,r) \leq c + c_d \int_0^T M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) \, dt + c_d \int_0^T \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,y)) \, dy dt \,. \qquad (41)
$$

Passing to the supremum for $0 < r < 2R$ we obtain

$$
\begin{aligned}
&\sup_{0\leq t\leq T} \sup_{0<r<2R} Q(t,x,r) \\
&\leq c + c_d \int_0^T M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) \, dt + c_d \int_0^T \sup_{0<r<2R} \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,y)) \, dy dt \,.
\end{aligned}
$$

Taking the L^p norm over $B_R(0)$ we get

$$
\begin{aligned}
A_p(R,\boldsymbol{X}) &= \left\| \sup_{0\leq t\leq T} \sup_{0<r<2R} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1 \right) \, dy \right\|_{L^p_x(B_R(0))} \\
&\leq c_{p,R} + c_d \left\| \int_0^T M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) \, dt \right\|_{L^p_x(B_R(0))} \\
&\quad + c_d \left\| \int_0^T \sup_{0<r<2R} \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,y)) \, dy dt \right\|_{L^p_x(B_R(0))} .
\end{aligned}
$$

Recalling Lemma 43 and the definition of the compressibility constant L, the first integral can be estimated with

$$
\begin{aligned}
c_d \int_0^T \left\| M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) \right\|_{L^p_x(B_R(0))} \, dt &\leq c_d L^{1/p} \int_0^T \left\| M_{\tilde{R}} D\boldsymbol{b}(t,x) \right\|_{L^p_x(B_{R+T\|\boldsymbol{b}\|_\infty}(0))} \, dt \\
&\leq c_{d,p} L^{1/p} \int_0^T \| D\boldsymbol{b}(t,x) \|_{L^p_x(B_{R+\tilde{R}+T\|\boldsymbol{b}\|_\infty}(0))} \, dt \,.
\end{aligned}
$$

The second integral can be estimated in a similar way with

$$
\begin{aligned}
& c_d \int_0^T \left\| \sup_{0<r<2R} \fint_{B_r(x)} \left[(M_{\tilde{R}} D\boldsymbol{b}) \circ (t, \boldsymbol{X}(t,\cdot)) \right](y)\, dy \right\|_{L^p_x(B_R(0))} dt \\
= \;& c_d \int_0^T \left\| M_{2R} \left[(M_{\tilde{R}} D\boldsymbol{b}) \circ (t, \boldsymbol{X}(t,\cdot)) \right](x) \right\|_{L^p_x(B_R(0))} dt \\
\leq \;& c_{d,p} \int_0^T \left\| \left[(M_{\tilde{R}} D\boldsymbol{b}) \circ (t, \boldsymbol{X}(t,\cdot)) \right](x) \right\|_{L^p_x(B_{3R}(0))} dt \\
= \;& c_{d,p} \int_0^T \left\| (M_{\tilde{R}} D\boldsymbol{b}) \circ (t, \boldsymbol{X}(t,x)) \right\|_{L^p_x(B_{3R}(0))} dt \\
\leq \;& c_{d,p} L^{1/p} \int_0^T \| M_{\tilde{R}} D\boldsymbol{b}(t,x) \|_{L^p_x(B_{3R+T\|\boldsymbol{b}\|_\infty}(0))} dt \\
\leq \;& c_{d,p} L^{1/p} \int_0^T \| D\boldsymbol{b}(t,x) \|_{L^p_x(B_{3R+T\|\boldsymbol{b}\|_\infty + \tilde{R}}(0))} dt \,.
\end{aligned}
$$

Then we obtain the desired estimate for $A_p(R, \boldsymbol{X})$. □

This estimate implies in an easy way a Lusin-type approximation of the flow with Lipschitz maps. As we observed before the proof of the proposition, the main point is that the estimate is now quantitative, and this will imply a precise control of the Lipschitz constant, in terms of the measure of the set we are "neglecting" (this explicit control was one of the open problems stated in [15]). This will be the key point in the proof of the compactness theorem.

Theorem 46 (Lipschitz estimates) *Let $\boldsymbol{b}$ be a vector field satisfying assumptions (A), (B) and (C) and denote by $\boldsymbol{X}$ its regular Lagrangian flow. Then, for every ϵ, $R > 0$, we can find a set $K \subset B_R(0)$ such that*

- $\mathscr{L}^d(B_R(0) \setminus K) \leq \epsilon$;
- *for any $0 \leq t \leq T$ we have*

$$
\mathrm{Lip}\left(\boldsymbol{X}(t,\cdot)|_K \right) \leq \exp \frac{c_d A_p(R, \boldsymbol{X})}{\epsilon^{1/p}} \,,
$$

with $A_p(R, \boldsymbol{X})$ satisfying the estimate of Proposition 45.

Proof. Fix $\epsilon > 0$ and $R > 0$. We apply Proposition 45 and equation (38) to obtain a constant

$$
M = M(\epsilon, p, A_p(R, \boldsymbol{X})) = \frac{A_p(R, \boldsymbol{X})}{\epsilon^{1/p}}
$$

and a set $K \subset B_R(0)$ with $\mathscr{L}^d(B_R(0) \setminus K) \leq \epsilon$ and

$$
\sup_{0 \leq t \leq T} \sup_{0<r<2R} \fint_{B_r(x)} \log \left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1 \right) dy \leq M \qquad \forall x \in K \,.
$$

This clearly means that

$$\fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{r}+1\right)dy \le M \qquad \text{for every } x\in K,\ t\in[0,T] \text{ and } r\in]0,2R[.$$

Now fix $x, y \in K$. Clearly $|x-y|<2R$. Set $r=|x-y|$ and compute

$$\begin{aligned}
&\log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{r}+1\right) = \fint_{B_r(x)\cap B_r(y)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{r}+1\right)dz \\
\le\ & \fint_{B_r(x)\cap B_r(y)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,z)|}{r}+1\right)+\log\left(\frac{|\boldsymbol{X}(t,y)-\boldsymbol{X}(t,z)|}{r}+1\right)dz \\
\le\ & c_d \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,z)|}{r}+1\right)dz + c_d \fint_{B_r(y)} \log\left(\frac{|\boldsymbol{X}(t,y)-\boldsymbol{X}(t,z)|}{r}+1\right)dz \\
\le\ & c_d M = \frac{c_d A_p(R,\boldsymbol{X})}{\epsilon^{1/p}}\,.
\end{aligned}$$

This implies that

$$|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)| \le \exp\left(\frac{c_d A_p(R,\boldsymbol{X})}{\epsilon^{1/p}}\right)|x-y| \qquad \text{for every } x,y\in K.$$

Therefore

$$\mathrm{Lip}(\boldsymbol{X}(t,\cdot)|_K) \ \le\ \exp\frac{c_d A_p(R,\boldsymbol{X})}{\epsilon^{1/p}}\,.$$

□

Recalling Theorem 39, the following corollary is an immediate consequence of the Lipschitz estimates.

Corollary 47 (Approximate differentiability of the flow) *Let $\boldsymbol{b}$ be a vector field satisfying assumptions (A), (B) and (C) and denote by $\boldsymbol{X}$ its regular Lagrangian flow. Then $\boldsymbol{X}(t,\cdot)$ is approximately differentiable $\mathscr{L}^d$-a.e. in $\mathbb{R}^d$, for every $t\in[0,T]$.*

Now we are going to present some new compactness results. The strategy of the proof is quite elementary: thanks to the Lipschitz estimates, the flows will be equi-continuous on large sets and this, together with the equi-boundedness in L^∞, will allow the use of Ascoli-Arzelà theorem "on big sets". From this, the compactness in L^1_{loc} easily follows.
The following theorem gives a partial answer to a conjecture raised by Alberto Bressan in [34]: in fact, Bressan's conjecture is the L^1 version of the L^p result that we are going to present. Presently this conjecture is still unsolved.

Theorem 48 (Compactness of the flow) *Let $\boldsymbol{b}_n$ be a sequence of smooth vector fields. Denote by $\boldsymbol{X}_n$ their regular Lagrangian flows and let L_n be the compressibility constant of the flow $\boldsymbol{X}_n$. Suppose that*

- *$|\boldsymbol{b}_n|$ are equi-bounded in $L^\infty([0,T]\times\mathbb{R}^d)$,*

- $|D_x \boldsymbol{b}_n|$ *are equi-bounded in* $L^1(L^p)$ *for some* $p > 1$*,*
- $\{L_n\}_n$ *is a bounded sequence (in* $\mathbb{R}$*).*

Then the sequence $\{\boldsymbol{X}_n\}_n$ *is relatively compact in* $L^1_{\mathrm{loc}}([0,T] \times \mathbb{R}^n)$*.*

Proof. Fix $R > 0$ and notice that, since the vector fields $\{\boldsymbol{b}_n\}_n$ are uniformly bounded in L^∞, the flows $\{\boldsymbol{X}_n\}_n$ are uniformly bounded in $L^\infty([0,T] \times B_R(0))$; denote by C_R a common bound in L^∞ for the flows. Applying Proposition 45 we obtain that the quantities $A_p(R, \boldsymbol{X}_n)$ are uniformly bounded with respect to n. Now, fix $j \in \mathbb{N}$ and apply Theorem 46 with $\epsilon = 1/j$ to obtain, for every $n \in \mathbb{N}$, a set $K_n \subset B_R(0)$ with $\mathscr{L}^d(B_R(0) \setminus K_n) \leq 1/j$ such that

$$\mathrm{Lip}(\boldsymbol{X}_n(t, \cdot)|_{K_n}) \quad \text{is uniformly bounded w.r.t. } n \text{ uniformly in } t.$$

Since $\frac{d}{dt}\boldsymbol{X}_n(t,x) = \boldsymbol{b}_n(t, \boldsymbol{X}_n(t,x))$, thanks to the equi-boundedness in L^∞ of the vector fields we also get that

$$\mathrm{Lip}(\boldsymbol{X}_n|_{[0,T]\times K_n}) \quad \text{is uniformly bounded w.r.t. } n.$$

Now, by applying some classical results on the extension of Lipschitz maps, we can extend every $\boldsymbol{X}_n|_{[0,T]\times K_n}$ to a map $\tilde{\boldsymbol{X}}_{n,j}$ defined on $[0,T] \times B_R(0)$ in such a way that

$$\mathrm{Lip}(\tilde{\boldsymbol{X}}_{n,j}|_{[0,T]\times B_R(0)}) \;\leq\; c_d \mathrm{Lip}(\boldsymbol{X}_n|_{[0,T]\times K_n}) \tag{42}$$

and

$$\|\tilde{\boldsymbol{X}}_{n,j}\|_{L^\infty([0,T]\times B_R(0))} \;\leq\; \|\boldsymbol{X}_n\|_{L^\infty([0,T]\times B_R(0))}\,. \tag{43}$$

Applying Ascoli–Arzelà theorem and a standard diagonal argument we can find a subsequence (not relabeled) such that for each j the sequence $\{\tilde{\boldsymbol{X}}_{n,j}\}_n$ converges uniformly (and hence strongly in $L^1([0,T] \times B_R(0))$) to a map $\tilde{\boldsymbol{X}}_{\infty,j}$.
Notice that

$$\|\tilde{\boldsymbol{X}}_{n,j} - \boldsymbol{X}_n\|_{L^1([0,T]\times B_R(0))} \leq \frac{2}{j} T \mathscr{L}^d(B_\rho(0)) C_R\,.$$

Next, for any given $\epsilon > 0$ select j such that

$$\frac{2}{j} T \mathscr{L}^d(B_R(0)) C_R \;\leq\; \epsilon/3\,,$$

and then $N > 0$ such that

$$\|\tilde{\boldsymbol{X}}_{i,j} - \tilde{\boldsymbol{X}}_{k,j}\|_{L^1([0,T]\times B_R(0))} \;\leq\; \epsilon/3 \qquad \text{for all } i,k > N\,.$$

Hence for every $i,k > N$ we get

$$\begin{aligned}\|\boldsymbol{X}_i - \boldsymbol{X}_k\|_{L^1([0,T]\times B_R(0))} \;&\leq\; \|\boldsymbol{X}_i - \tilde{\boldsymbol{X}}_{i,j}\|_{L^1([0,T]\times B_R(0))} + \|\tilde{\boldsymbol{X}}_{i,j} - \tilde{\boldsymbol{X}}_{k,j}\|_{L^1([0,T]\times B_R(0))} \\ &\quad + \|\boldsymbol{X}_k - \tilde{\boldsymbol{X}}_{k,j}\|_{L^1([0,T]\times B_R(0))} \;\leq\; \epsilon\,.\end{aligned}$$

Hence $\{\boldsymbol{X}_n\}_n$ is a Cauchy sequence in $L^1([0,T] \times B_R(0))$. A second diagonal argument yields a subsequence, not relabeled, which is a Cauchy sequence in $L^1([0,T] \times B_l(0))$ for every $l \in \mathbb{N}$. □

Remark 49 (Existence of the regular Lagrangian flow) As a corollary we obtain a new proof of the existence of a regular Lagrangian flow associated to a bounded vector field $\boldsymbol{b}$ satisfying assumptions (A), (B) and (C). Indeed, approximating $\boldsymbol{b}$ by convolution with a positive convolution kernel, we get a sequence which satisfies the assumptions of the previous theorem, hence in the limit we get a flow associated to $\boldsymbol{b}$, thanks to the compactness in the strong topology. An analogous remark applies to Theorem 52.

With similar techniques it is also possible to show compactness in the $W^{1,1}$ case, under the assumption that the maximal functions of the derivatives of the vector fields $\boldsymbol{b}_n$ are uniformly bounded in $L^1(L^1)$. The strategy is slightly different, since we are not able to show Lipschitz estimates under this weaker assumption (notice that the proof of Proposition 45 requires a control of $M[MD\boldsymbol{b}]$). We start defining an integral quantity similar to the previous one, but without the supremum with respect to the radius r. For $R > 0$ and $0 < r < R/2$ fixed we set

$$a(r, R, \boldsymbol{X}) = \int_{B_R(0)} \sup_{0\le t\le T} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1\right) dydx\,.$$

We first give a quantitative estimate for the quantity $a(r, R, \boldsymbol{X})$, similar to the previous one for $A_p(R, \boldsymbol{X})$.

Proposition 50 *Let $\boldsymbol{b}$ be a vector field satisfying the assumptions (A) and (C) and such that*

$$M_\lambda D\boldsymbol{b} \in L^1(L^1) \qquad \textit{for every } \lambda > 0.$$

Denote by $\boldsymbol{X}$ its regular Lagrangian flow and let L be the compressibility constant of the flow. Then we have

$$a(r, R, \boldsymbol{X}) \le C\left(R, L, \|M_{\tilde{R}} D_x \boldsymbol{b}\|_{L^1(L^1)}\right)\,,$$

where $\tilde{R} = 3R/2 + 2T\|\boldsymbol{b}\|_\infty$.

Proof. We start as in the proof of Proposition 45, obtaining inequality (41) (but this time it is sufficient to set $\tilde{R} = 3R/2 + 2T\|\boldsymbol{b}\|_\infty$). Integrating with respect to x over $B_R(0)$, we obtain

$$\begin{aligned}
a(r, R, \boldsymbol{X}) &= \int_{B_R(0)} \sup_{0\le t\le T} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1\right) dydx \\
&\le c_R + c_d \int_{B_R(0)} \int_0^T M_{\tilde{R}} D\boldsymbol{b}(t, \boldsymbol{X}(t,x))\, dtdx \\
&\quad + c_d \int_{B_R(0)} \int_0^T \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t, \boldsymbol{X}(t,y))\, dydtdx\,.
\end{aligned}$$

As in the previous computations, the first integral can be estimated with

$$c_d L \left\| M_{\tilde{R}} D\boldsymbol{b} \right\|_{L^1([0,T];L^1(B_{R+T\|\boldsymbol{b}\|_\infty}(0)))}\,,$$

but this time we cannot bound the norm of the maximal function with the norm of the derivative. To estimate the last integral we compute

$$\begin{aligned}
& c_d \int_{B_R(0)} \int_0^T \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t, \boldsymbol{X}(t,y))\, dydtdx \\
= \;& c_d \int_{B_R(0)} \int_0^T \fint_{B_r(0)} M_{\tilde{R}} D\boldsymbol{b}(t, \boldsymbol{X}(t,x+z))\, dzdtdx \\
\leq \;& c_d \fint_{B_r(0)} \int_0^T \int_{B_R(0)} M_{\tilde{R}} D\boldsymbol{b}(t, \boldsymbol{X}(t,x+z))\, dxdtdz \\
\leq \;& c_d \fint_{B_r(0)} \int_0^T L \int_{B_{3R/2+T\|\boldsymbol{b}\|_\infty}(0)} M_{\tilde{R}} D\boldsymbol{b}(t,w)\, dwdtdx \\
= \;& c_d L \| M_{\tilde{R}} D\boldsymbol{b}\|_{L^1([0,T];L^1(B_{3R/2+T\|\boldsymbol{b}\|_\infty}(0)))}\,.
\end{aligned}$$

This concludes the proof of the estimate for $a(r, R, \boldsymbol{X})$. □

We recall the following classical criterion for strong compactness in L^p, since it will be used in the proof of the compactness theorem. For the proof of the lemma we refer for example to [35].

Lemma 51 (Riesz-Fréchet-Kolmogorov compactness criterion) *Let $\mathscr{F}$ be a bounded subset of $L^p(\mathbb{R}^N)$ for some $1 \leq p < \infty$. Suppose that*

$$\lim_{|h|\to 0} \|f(\cdot - h) - f\|_{L^p} = 0 \qquad \textit{uniformly in } f \in \mathscr{F}\,.$$

Then $\mathscr{F}$ is relatively compact in $L^p_{\mathrm{loc}}(\mathbb{R}^N)$.

Theorem 52 *Let $\boldsymbol{b}_n$ be a sequence of smooth vector fields. Denote by $\boldsymbol{X}_n$ their regular Lagrangian flows and let L_n be the compressibility constant of the flow $\boldsymbol{X}_n$. Suppose that*

- *$|\boldsymbol{b}_n|$ are equi-bounded in $L^\infty([0,T] \times \mathbb{R}^d)$,*
- *$M_\lambda D_x \boldsymbol{b}_n$ are equi-bounded in $L^1(L^1)$ for every $\lambda > 0$,*
- *$\{L_n\}_n$ is a bounded sequence (in $\mathbb{R}$).*

Then the sequence $\{\boldsymbol{X}_n\}_n$ is relatively compact in $L^1_{\mathrm{loc}}([0,T] \times \mathbb{R}^d)$.

Proof. We apply Proposition 50 to obtain that, under the assumptions of the Theorem, the quantities $a(r, R, \boldsymbol{X}_n)$ are uniformly bounded with respect to n. Now observe that, for $0 \leq z \leq \bar{R}$ (with $\bar{R} = 3R/2 + 2T\|\boldsymbol{b}\|_\infty$ as in Proposition 50), thanks to the concavity of the logarithm we have

$$\log\left(\frac{z}{r} + 1\right) \geq \frac{\log\left(\frac{\bar{R}}{r} + 1\right)}{\bar{R}} z\,.$$

Since $|\boldsymbol{X}_n(t,x) - \boldsymbol{X}_n(t,y)| \leq \bar{R}$ this implies that

$$\begin{aligned} &\int_{B_R(0)} \sup_{0\leq t\leq T} \fint_{B_r(x)} |\boldsymbol{X}_n(t,x) - \boldsymbol{X}_n(t,y)|\, dydx \\ \leq\ & \frac{\bar{R}}{\log\left(\frac{\bar{R}}{r}+1\right)} C\left(R, L_n, \|M_{\bar{R}} D_x \boldsymbol{b}_n\|_{L^1(L^1)}\right) \ \leq\ g(r)\,, \end{aligned}$$

where the function $g(r)$ does not depend on n and satisfies $g(r) \downarrow 0$ for $r \downarrow 0$. Changing the integration order this implies

$$\fint_{B_r(0)} \int_{B_R(0)} |\boldsymbol{X}_n(t,x) - \boldsymbol{X}_n(t,x+z)|\, dxdz \leq g(r)\,,$$

uniformly with respect to t and n.

Now notice the following elementary fact. There exists a dimensional constant $\alpha_d > 0$ with the following property: if $A \subset B_1(0)$ is a measurable set with $\mathscr{L}^d(B_1(0) \setminus A) \leq \alpha_d$, then $A + A \supset B_{1/2}(0)$. Then fix α_d as above and apply Chebyshev inequality for every n to obtain, for every $0 < r < R/2$, a measurable set $K_{r,n} \subset B_r(0)$ with $\mathscr{L}^d(B_r(0) \setminus K_{r,n}) \leq \alpha_d \mathscr{L}^d(B_r(0))$ and

$$\int_{B_R(0)} |\boldsymbol{X}_n(t,x+z) - \boldsymbol{X}_n(t,x)|\, dx \leq \frac{g(r)}{\alpha_d} \qquad \text{for every } z \in K_{r,n}.$$

For such a set $K_{r,n}$, thanks to the previous remark, we have that $K_{r,n} + K_{r,n} \supset B_{r/2}(0)$. Now let $v \in B_{r/2}(0)$ be arbitrary. For every n we can write $v = z_{1,n} + z_{2,n}$ with $z_{1,n}, z_{2,n} \in K_{r,n}$. We can estimate the increment in the spatial directions as follows:

$$\begin{aligned} &\int_{B_{R/2}(0)} |\boldsymbol{X}_n(t,x+v) - \boldsymbol{X}_n(t,x)|\, dx \\ =\ & \int_{B_{R/2}(0)} |\boldsymbol{X}_n(t,x+z_{1,n}+z_{2,n}) - \boldsymbol{X}_n(t,x)|\, dx \\ \leq\ & \int_{B_{R/2}(0)} |\boldsymbol{X}_n(t,x+z_{1,n}+z_{2,n}) - \boldsymbol{X}_n(t,x+z_{1,n})| + |\boldsymbol{X}_n(t,x+z_{1,n}) - \boldsymbol{X}_n(t,x)|\, dx \\ \leq\ & \int_{B_R(0)} |\boldsymbol{X}_n(t,y+z_{2,n}) - \boldsymbol{X}_n(t,y)|\, dy + \int_{B_R(0)} |\boldsymbol{X}_n(t,x+z_{1,n}) - \boldsymbol{X}_n(t,x)|\, dx \ \leq\ \frac{2g(r)}{\alpha_d}\,. \end{aligned}$$

Now notice that, by definition of regular Lagrangian flow, we have

$$\frac{d\boldsymbol{X}_n}{dt}(t,x) = \boldsymbol{b}_n(t, \boldsymbol{X}_n(t,x))\,.$$

Then we can estimate the increment in the time direction in the following way

$$|\boldsymbol{X}_n(t+h,x) - \boldsymbol{X}_n(t,x)| \leq \int_0^h \left|\frac{d\boldsymbol{X}_n}{dt}(t+s,x)\right| ds = \int_0^h |\boldsymbol{b}_n(t+s, \boldsymbol{X}_n(t+s,x))|\, ds \leq h\|\boldsymbol{b}_n\|_\infty\,.$$

Combining these two informations, for $(t_0, t_1) \subset\subset [0,T]$, $R > 0$, $v \in B_{r/2}(0)$ and $h > 0$ sufficently small we can estimate

$$\begin{aligned}
&\int_{t_0}^{t_1}\int_{B_{R/2}(0)} |\boldsymbol{X}_n(t+h, x+v) - \boldsymbol{X}_n(t,x)|\, dxdt \\
\leq\ & \int_{t_0}^{t_1}\int_{B_{R/2}(0)} |\boldsymbol{X}_n(t+h, x+v) - \boldsymbol{X}_n(t+h,x)| + |\boldsymbol{X}_n(t+h, x) - \boldsymbol{X}_n(t,x)|\, dxdt \\
\leq\ & T\frac{2g(r)}{\alpha_d} + \int_{t_0}^{t_1}\int_{B_{R/2}(0)} h\|\boldsymbol{b}_n\|_\infty\, dxdt \ \leq\ T\frac{2g(r)}{\alpha_d} + c_d T R^d h \|\boldsymbol{b}_n\|\,.
\end{aligned}$$

The thesis follows applying the Riesz-Fréchet-Kolmogorov compactness criterion, recalling that $\boldsymbol{b}_n$ are uniformly bounded in L^∞. $\square$

We conclude this section showing another result obtained in [48] with techniques which are very similar to the ones described so far. It is a result of quantitative stability for regular Lagrangian flows: it improves the stability result given in Theorem 29 in the sense that it gives a rate of convergence of the flows in terms of convergence of the vector fields.

Theorem 53 (Quantitative stability) *Let $\boldsymbol{b}$ and $\tilde{\boldsymbol{b}}$ be vector fields satisfying assumptions (A), (B) and (C). Denote by $\boldsymbol{X}$ and $\tilde{\boldsymbol{X}}$ the respective regular Lagrangian flows and let L and $\tilde{L}$ be the compressibility constants of the two flows. Then, for every time $\tau \in [0,T]$, we have*

$$\|\boldsymbol{X}(\tau,\cdot) - \tilde{\boldsymbol{X}}(\tau,\cdot)\|_{L^1(B_r(0))} \leq C \left|\log\left(\|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_{L^1([0,\tau]\times B_R(0))}\right)\right|^{-1},$$

where $R = r + T\|\boldsymbol{b}\|_\infty$ and the constant C only depends on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$.

Proof. Set $\delta := \|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_{L^1([0,T]\times B_R(0))}$ and consider the function

$$g(t) \ := \ \int_{B_r(0)} \log\left(\frac{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)|}{\delta} + 1\right) dx\,.$$

Clearly $g(0) = 0$ and after some standard computations we get

$$\begin{aligned}
g'(t) \ &\leq\ \int_{B_r(0)} \left|\frac{d\boldsymbol{X}(t,x)}{dt} - \frac{d\tilde{\boldsymbol{X}}(t,x)}{dt}\right| \left(|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta\right)^{-1} dx \\
&=\ \int_{B_r(0)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x)) - \tilde{\boldsymbol{b}}(t,\tilde{\boldsymbol{X}}(t,x))|}{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta}\, dx \\
&\leq\ \frac{1}{\delta}\int_{B_r(0)} |\boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x)) - \tilde{\boldsymbol{b}}(t,\tilde{\boldsymbol{X}}(t,x))|\, dx \\
&\quad + \int_{B_r(0)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x)) - \boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x))|}{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta}\, dx\,.
\end{aligned} \tag{44}$$

We set $\tilde{R} = 2r + T(\|\boldsymbol{b}\|_\infty + \|\tilde{\boldsymbol{b}}\|_\infty)$ and we apply Lemma 44 to estimate the last integral as follows:

$$\int_{B_r(0)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x)) - \boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x))|}{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta}\, dx \le c_d \int_{B_r(0)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) + M_{\tilde{R}} D\boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x))\, dx\,.$$

Inserting this estimate in (44), setting $\tilde{r} = r + T\max\{\|\boldsymbol{b}\|_\infty, \|\tilde{\boldsymbol{b}}\|_\infty\}$, changing variables in the integrals and using Lemma 43 we get

$$\begin{aligned}
g'(t) &\le \frac{\tilde{L}}{\delta}\int_{B_{r+T\|\tilde{\boldsymbol{b}}\|_\infty}(0)} |\boldsymbol{b}(t,y) - \tilde{\boldsymbol{b}}(t,y)|\, dy + \big(\tilde{L} + L\big)\int_{B_{\tilde{r}}(0)} M_{\tilde{R}} D\boldsymbol{b}(t,y)\, dy \\
&\le \frac{\tilde{L}}{\delta}\int_{B_{r+T\|\tilde{\boldsymbol{b}}\|_\infty}(0)} |\boldsymbol{b}(t,y) - \tilde{\boldsymbol{b}}(t,y)|\, dy + c_d \tilde{r}^{n-n/p}\big(\tilde{L} + L\big)\|M_{\tilde{R}} D\boldsymbol{b}(t,\cdot)\|_{L^p_x} \\
&\le \frac{\tilde{L}}{\delta}\int_{B_{r+T\|\tilde{\boldsymbol{b}}\|_\infty}(0)} |\boldsymbol{b}(t,y) - \tilde{\boldsymbol{b}}(t,y)|\, dy + c_{d,p} \tilde{r}^{n-n/p}\big(\tilde{L} + L\big)\|D\boldsymbol{b}(t,\cdot)\|_{L^p_x}\,.
\end{aligned}$$

For any $\tau \in [0,T]$, integrating the last inequality between 0 and τ we get

$$g(\tau) \;=\; \int_{B_r(0)} \log\left(\frac{|\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|}{\delta} + 1\right) dx \;\le\; C_1\,, \tag{45}$$

where the constant C_1 depends on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$.
Next we fix a second parameter $\eta > 0$ to be chosen later. Using Chebyshev inequality we find a measurable set $K \subset B_r(0)$ such that $\mathscr{L}^d(B_r(0)\setminus K) \le \eta$ and

$$\log\left(\frac{|\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|}{\delta} + 1\right) \;\le\; \frac{C_1}{\eta} \qquad \text{for } x \in K.$$

Therefore we can estimate

$$\begin{aligned}
&\int_{B_r(0)} |\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|\, dx \\
\le\;& \eta\left(\|\boldsymbol{X}(\tau,\cdot)\|_{L^\infty(B_r(0))} + \|\tilde{\boldsymbol{X}}(\tau,\cdot)\|_{L^\infty(B_r(0))}\right) + \int_K |\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|\, dx \\
\le\;& \eta C_2 + c_d r^n \delta\left(\exp(C_1/\eta)\right) \;\le\; C_3\left(\eta + \delta\exp(C_1/\eta)\right)\,,
\end{aligned} \tag{46}$$

with C_1, C_2 and C_3 which depend only on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$. Without loss of generality we can assume $\delta < 1$. Setting $\eta = 2C_1|\log\delta|^{-1} = 2C_1(-\log\delta)^{-1}$, we have $\exp(C_1/\eta) = \delta^{-1/2}$. Thus we conclude

$$\int_{B_r(0)} |\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|\, dx \;\le\; C_3\left(2C_1|\log\delta|^{-1} + \delta^{1/2}\right) \;\le\; C|\log\delta|^{-1}\,, \tag{47}$$

where C depends only on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$. This completes the proof.

□

Remark 54 (Uniqueness of the regular Lagrangian flow) We observe that the previous theorem also gives a new proof of the uniqueness of the regular Lagrangian flow associated to a vector field $\boldsymbol{b}$ which satisfies assumptions (A), (B) and (C).

9 Bibliographical notes and open problems

Section 3. The material contained in this section is classical. Good references are [64], Chapter 8 of [14], [27] and [61]. For the proof of the area formula, see for instance [13], [67], [68]. The proof of the second local variant, under the stronger assumption $\int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t| \, d\mu_t dt < \infty$, is given in Proposition 8.1.8 of [14]. The same proof works under the weaker assumption (6).

Section 4. Many ideas of this section, and in particular the idea of looking at measures in the space of continuous maps to characterize the flow and prove its stability, are borrowed from [6], dealing with BV vector fields. Later on, the arguments have been put in a more general form, independent of the specific class of vector fields under consideration, in [7]. Here we present the same version of [8].

The idea of a probabilistic representation is of course classical, and appears in many contexts (particularly for equations of diffusion type); to our knowledge the first reference in the context of conservation laws and fluid mechanics is [30], where a similar approach is proposed for building generalized geodesics in the space $\mathbf{G}$ of measure-preserving diffeomorphisms; this is related to Arnold's interpretation of the incompressible Euler equation as a geodesic in $\mathbf{G}$ (see also [31], [32], [33]): in this case the compact (but neither metrizable, nor separable) space $X^{[0,T]}$, with $X \subset \mathbb{R}^d$ compact, has been considered.

This approach is by now a familiar one also in optimal transport theory, where transport maps and transference plans can be thought in a natural way as measures in the space of minimizing geodesics [85], and in the so called irrigation problems, a nice variant of the optimal transport problem [23]. See also [20] for a similar approach within Mather's theory. The lecture notes [94] (see also the Appendix of [78]) contain, among several other things, a comprehensive treatment of the topic of measures in the space of action-minimizing curves, including at the same time the optimal transport and the dynamical systems case (this unified treatment was inspired by [22]). Another related reference is [58].

The superposition principle is proved, under the weaker assumption $\int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t|^p \, d\mu_t dt < +\infty$ for some $p > 1$, in Theorem 8.2.1 of [14], see also [79] for the extension to the case $p = 1$ and to the non-homogeneous continuity equation. Very closely related results, relative to the representation of a vector field as the superposition of "elementary" vector fields associated to curves, appear in [86], [20].

Section 5. The definition of renormalized solution and the strong convergence of commutators are entirely borrowed from [61]. See also [62] for the relevance of this concept in connection with the existence theory for Boltzmann equation. The proof of the comparison principle assuming only an $L^1(L^1_{\mathrm{loc}})$ bound (instead of an $L^1(L^\infty)$ one, as in [61], [6]) on the divergence was suggested to us by G.Savaré. See also [40] for a proof, using radial convolution kernels, of the renormalization property for vector fields satisfying $D_i \boldsymbol{b}^j + D_j \boldsymbol{b}^i \in L^1_{\mathrm{loc}}$.

No general existence result for Sobolev (or even BV) vector fields seems to be known in the

infinite-dimensional case: the only reference we are aware of is [24], dealing with vector-fields having an exponentially integrable derivative, extending previous results in [49], [50], [51]. Also the investigation of non-Euclidean geometries, *e.g.* Carnot groups and horizontal vector fields, could provide interesting results.
Finally, notice that the theory has a natural invariance, namely if $\boldsymbol{X}$ is a flow relative to $\boldsymbol{b}$, then $\boldsymbol{X}$ is a flow relative to $\tilde{\boldsymbol{b}}$ whenever $\{\tilde{\boldsymbol{b}} \neq \boldsymbol{b}\}$ is $\mathscr{L}^{1+d}$-negligible in $(0,T)\times\mathbb{R}^d$. So a natural question is whether the uniqueness "in the selection sense" might be enforced by choosing a canonical representative $\tilde{\boldsymbol{b}}$ in the equivalence class of $\boldsymbol{b}$: in other words we may think that, for a suitable choice of $\tilde{\boldsymbol{b}}$, the ODE $\dot\gamma(t) = \tilde{\boldsymbol{b}}_t(\gamma(t))$ has a unique absolutely continuous solution starting from x for $\mathscr{L}^d$-a.e. x.
Concerning the case of BV vector fields, the main idea of the proof of Theorem 33, *i.e.* the adaptation of the convolution kernel to the local behaviour of the vector field, has been used at various level of generality in [25], [77], [45] (see also [41], [42] for related results independent of this technique), until the general result obtained in [6].
The optimal regularity condition on $\boldsymbol{b}$ ensuring the renormalization property, and therefore the validity of the comparison principle in $\mathscr{L}_{\boldsymbol{b}}$, is still not known. New results, both in the Sobolev and in the BV framework, are presented in [10], [73], [74].
In [12] we investigate in particular the possibility to prove the renormalization property for nearly incompressible $BV_{\mathrm{loc}}\cap L^\infty$ fields $\boldsymbol{b}$: they are defined by the property that there exists a positive function ρ, with $\ln\rho\in L^\infty$, such that the space-time field $(\rho,\rho\boldsymbol{b})$ is divergence free. As in the case of the Keyfitz-Kranzer system, the existence a function ρ with this property seems to be a natural replacement of the condition $D_x\cdot\boldsymbol{b}\in L^\infty$ (and is actually implied by it); as explained in [9], a proof of the renormalization property in this context would lead to a proof of a conjecture, due to Bressan, on the compactness of flows associated to a sequence of vector fields bounded in $BV_{t,x}$.

Section 6. The material of this section is entirely taken from [17]. See Chapter 3 of [68] for a deep study of approximate differentiability and calculus with the class of approximately differentiable maps.

Section 7. Here we have presented the main result in [73].

Section 8. The first progress in the direction of proving approximate differentiability of the flow $\boldsymbol{X}(t,x)$ with respect to x has been achieved in [15]. Later on, these results have been substantially improved in [47], and for this reason we have chosen to present only the latter results and proofs in this section.

References

[1] M.Aizenman: *On vector fields as generators of flows: a counterexample to Nelson's conjecture.* Ann. Math., **107** (1978), 287–296.

[2] G.Alberti: *Rank-one properties for derivatives of functions with bounded variation.* Proc. Roy. Soc. Edinburgh Sect. A, **123** (1993), 239–274.

[3] G.Alberti & L.Ambrosio: *A geometric approach to monotone functions in $\mathbb{R}^n$.* Math. Z., **230** (1999), 259–316.

[4] G.Alberti & S.Müller: *A new approach to variational problems with multiple scales.* Comm. Pure Appl. Math., **54** (2001), 761–825.

[5] F.J.Almgren: *The theory of varifolds – A variational calculus in the large,* Princeton University Press, 1972.

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

[7] L.Ambrosio: *Lecture notes on transport equation and Cauchy problem for BV vector fields and applications.* Preprint, 2004 (available at http://cvgmt.sns.it).

[8] L.Ambrosio: *Lecture notes on transport equation and Cauchy problem for non-smooth vector fields and applications.* Preprint, 2005 (available at http://cvgmt.sns.it).

[9] L.Ambrosio, F.Bouchut & C.De Lellis: *Well-posedness for a class of hyperbolic systems of conservation laws in several space dimensions.* Comm. PDE, **29** (2004), 1635–1651.

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

[11] L.Ambrosio & C.De Lellis: *Existence of solutions for a class of hyperbolic systems of conservation laws in several space dimensions.* International Mathematical Research Notices, **41** (2003), 2205–2220.

[12] L.Ambrosio, C.De Lellis & J.Malý: *On the chain rule for the divergence of BV like vector fields: applications, partial results, open problems.* To appear in the forthcoming book by the AMS series in contemporary mathematics "Perspectives in Nonlinear Partial Differential Equations: in honor of Haim Brezis" (available at http://cvgmt.sns.it).

[13] L.Ambrosio, N.Fusco & D.Pallara: *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, 2000.

[14] L.Ambrosio, N.Gigli & G.Savaré: *Gradient flows in metric spaces and in the Wasserstein space of probability measures.* Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.

[15] L.Ambrosio, M.Lecumberry & S.Maniglia: *Lipschitz regularity and approximate differentiability of the DiPerna-Lions flow.* Rendiconti del Seminario Fisico Matematico di Padova, **114** (2005), 29–50.

[16] L.Ambrosio, S.Lisini & G.Savaré: *Stability of flows associated to gradient vector fields and convergence of iterated transport maps.* Manuscripta Math., **121** (2006), 1–50.

[17] L.Ambrosio & J.Malý: *Very weak notions of differentiability.* Preprint, 2005 (available at http://cvgmt.sns.it).

[18] E.J.Balder: *New fundamentals of Young measure convergence.* CRC Res. Notes in Math. **411**, 2001.

[19] J.Ball & R.James: *Fine phase mixtures as minimizers of energy.* Arch. Rat. Mech. Anal., **100** (1987), 13–52.

[20] V.Bangert: *Minimal measures and minimizing closed normal one-currents.* Geom. funct. anal., **9** (1999), 413–427.

[21] J.-D.Benamou & Y.Brenier: *Weak solutions for the semigeostrophic equation formulated as a couples Monge-Ampere transport problem.* SIAM J. Appl. Math., **58** (1998), 1450–1461.

[22] P.Bernard & B.Buffoni: *Optimal mass transportation and Mather theory.* J. Eur. Math. Soc. (JEMS), **9** (2007), 85–121.

[23] M.Bernot, V.Caselles & J.M.Morel: *Traffic plans.* Publ. Mat., **49** (2005), 417–451.

[24] V.Bogachev & E.M.Wolf: *Absolutely continuous flows generated by Sobolev class vector fields in finite and infinite dimensions.* J. Funct. Anal., **167** (1999), 1–68.

[25] F.Bouchut: *Renormalized solutions to the Vlasov equation with coefficients of bounded variation.* Arch. Rational Mech. Anal., **157** (2001), 75–90.

[26] F.Bouchut & G.Crippa: *Uniqueness, Renormalization, and Smooth Approximations for Linear Transport Equations.* SIAM J. Math. Anal., **38** (2006), 1316–1328.

[27] F.Bouchut, F.Golse & M.Pulvirenti: *Kinetic equations and asymptotic theory.* Series in Appl. Math., Gauthiers-Villars, 2000.

[28] F.Bouchut & F.James: *One dimensional transport equation with discontinuous coefficients.* Nonlinear Analysis, **32** (1998), 891–933.

[29] F.Bouchut, F.James & S.Mancini: *Uniqueness and weak stability for multi-dimensional transport equations with one-sided Lipschitz coefficients.* Annali Scuola Normale Superiore, Ser. 5, **4** (2005), 1–25.

[30] Y.Brenier: *The least action principle and the related concept of generalized flows for incompressible perfect fluids.* J. Amer. Mat. Soc., **2** (1989), 225–255.

[31] Y.Brenier: *The dual least action problem for an ideal, incompressible fluid.* Arch. Rational Mech. Anal., **122** (1993), 323–351.

[32] Y.Brenier: *A homogenized model for vortex sheets.* Arch. Rational Mech. Anal., **138** (1997), 319–353.

[33] Y.Brenier: *Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations.* Comm. Pure Appl. Math., **52** (1999), 411–452.

[34] A.Bressan: *An ill posed Cauchy problem for a hyperbolic system in two space dimensions.* Rend. Sem. Mat. Univ. Padova, **110** (2003), 103–117.

[35] H.Brezis: *Analyse fonctionnelle. Théorie et applications.* Masson, Paris, 1983.

[36] L.A.Caffarelli: *Some regularity properties of solutions of Monge Ampère equation*, Comm. Pure Appl. Math., **44** (1991), 965–969.

[37] L.A.Caffarelli: *Boundary regularity of maps with convex potentials*, Comm. Pure Appl. Math., **45** (1992), 1141–1151.

[38] L.A.Caffarelli: *The regularity of mappings with a convex potential.* J. Amer. Math. Soc., **5** (1992), 99–104.

[39] L.A.Caffarelli: *Boundary regularity of maps with convex potentials.*, Ann. of Math., **144** (1996), 453–496.

[40] I.Capuzzo Dolcetta & B.Perthame: *On some analogy between different approaches to first order PDE's with nonsmooth coefficients.* Adv. Math. Sci Appl., **6** (1996), 689–703.

[41] A.Cellina: *On uniqueness almost everywhere for monotonic differential inclusions.* Nonlinear Analysis, TMA, **25** (1995), 899–903.

[42] A.Cellina & M.Vornicescu: *On gradient flows.* Journal of Differential Equations, **145** (1998), 489–501.

[43] F.Colombini, G.Crippa & J.Rauch: *A note on two dimensional transport with bounded divergence.* Comm. Partial Differential Equations, **31** (2006), 1109–1115.

[44] F.Colombini & N.Lerner: *Uniqueness of continuous solutions for BV vector fields.* Duke Math. J., **111** (2002), 357–384.

[45] F.Colombini & N.Lerner: *Uniqueness of L^∞ solutions for a class of conormal BV vector fields.* Contemp. Math. **368** (2005), 133–156.

[46] F.Colombini, T. Luo & J.Rauch: *Nearly Lipschitzean diverge free transport propagates neither continuity nor BV regularity.* Commun. Math. Sci., **2** (2004), 207–212.

[47] G.Crippa & C.De Lellis: *Oscillatory solutions to transport equations.* Indiana Univ. Math. J., **55** (2006), 1–13.

[48] G.Crippa & C.De Lellis: *Estimates and regularity results for the DiPerna–Lions flow.* Preprint, 2006 (available at http://cvgmt.sns.it). Accepted by J. Reine Angew. Math.

[49] A.B.Cruzeiro: *Équations différentielles ordinaires: non explosion et mesures quasi-invariantes.* J. Funct. Anal., **54** (1983), 193–205.

[50] A.B.Cruzeiro: *Équations différentielles sur l'espace de Wiener et formules de Cameron-Martin non linéaires.* J. Funct. Anal., **54** (1983), 206–227.

[51] A.B.Cruzeiro: *Unicité de solutions d'équations différentielles sur l'espace de Wiener.* J. Funct. Anal., **58** (1984), 335–347.

[52] M.Cullen: *On the accuracy of the semi-geostrophic approximation.* Quart. J. Roy. Metereol. Soc., **126** (2000), 1099–1115.

[53] M.Cullen & M.Feldman: *Lagrangian solutions of semigeostrophic equations in physical space.* SIAM J. Math. Anal., **37** (2006), 1371–1395.

[54] M.Cullen & W.Gangbo: *A variational approach for the 2-dimensional semi-geostrophic shallow water equations.* Arch. Rational Mech. Anal., **156** (2001), 241–273.

[55] C.Dafermos: *Hyperbolic conservation laws in continuum physics.* Springer Verlag, 2000.

[56] N.G.de Bruijn: *On almost additive functions.* Colloq. Math. **15** (1966), 59–63.

[57] C.De Lellis: *Blow-up of the BV norm in the multidimensional Keyfitz and Kranzer system.* Duke Math. J., **127** (2004), 313–339.

[58] L.De Pascale, M.S.Gelli & L.Granieri: *Minimal measures, one-dimensional currents and the Monge-Kantorovich problem.* Calc. Var. Partial Differential Equations, **27** (2006), 1–23.

[59] N.Depauw: *Non unicité des solutions bornées pour un champ de vecteurs BV en dehors d'un hyperplan.* C.R. Math. Sci. Acad. Paris, **337** (2003), 249–252.

[60] R.J.DiPerna: *Measure-valued solutions to conservation laws.* Arch. Rational Mech. Anal., **88** (1985), 223–270.

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

[62] R.J.DiPerna & P.L.Lions: *On the Cauchy problem for the Boltzmann equation: global existence and weak stability.* Ann. of Math., **130** (1989), 312–366.

[63] L.C.Evans: *Partial Differential Equations and Monge–Kantorovich Mass Transfer.* Current Developments in Mathematics (1997), 65–126.

[64] L.C.Evans: *Partial Differential Equations.* Graduate studies in Mathematics, **19** (1998), American Mathematical Society.

[65] L.C.Evans & W.Gangbo: *Differential equations methods for the Monge-Kantorovich mass transfer problem.* Memoirs AMS, **653**, 1999.

[66] L.C.Evans, W.Gangbo & O.Savin: *Diffeomorphisms and nonlinear heat flows.* SIAM J. Math. Anal., **37** (2005), 737–751.

[67] L.C.Evans & R.F.Gariepy: *Lecture notes on measure theory and fine properties of functions,* CRC Press, 1992.

[68] H.Federer: *Geometric measure theory*, Springer, 1969.

[69] M.Hauray: *On Liouville transport equation with potential in $BV_{\rm loc}$*. Comm. Partial Differential Equations, **29** (2004), 207–217.

[70] M.Hauray: *On two-dimensional Hamiltonian transport equations with $L^p_{\rm loc}$ coefficients.* Ann. IHP Nonlinear Anal. Non Linéaire, **20** (2003), 625–644.

[71] W.B.Jurkat: *On Cauchy's functional equation.* Proc. Amer. Math. Soc., **16** (1965), 683–686.

[72] B.L.Keyfitz & H.C.Kranzer: *A system of nonstrictly hyperbolic conservation laws arising in elasticity theory.* Arch. Rational Mech. Anal. **1980**, *72*, 219–241.

[73] C.Le Bris & P.L.Lions: *Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications.* Annali di Matematica, **183** (2003), 97–130.

[74] N.Lerner: *Transport equations with partially BV velocities.* Ann. Sc. Norm. Super. Pisa Cl. Sci., **3** (2004), 681–703.

[75] P.L.Lions: *Mathematical topics in fluid mechanics, Vol. I: incompressible models.* Oxford Lecture Series in Mathematics and its applications, **3** (1996), Oxford University Press.

[76] P.L.Lions: *Mathematical topics in fluid mechanics, Vol. II: compressible models.* Oxford Lecture Series in Mathematics and its applications, **10** (1998), Oxford University Press.

[77] P.L.Lions: *Sur les équations différentielles ordinaires et les équations de transport.* C. R. Acad. Sci. Paris Sér. I, **326** (1998), 833–838.

[78] J.Lott & C.Villani: *Weak curvature conditions and functional inequalities.* J. Funct. Anal., in press.

[79] S.Maniglia: *Probabilistic representation and uniqueness results for measure-valued solutions of transport equations.* Preprint, 2005 (available at http://cvgmt.sns.it). Accepted by J. Math. Pures Appl.

[80] J.N.Mather: *Minimal measures.* Comment. Math. Helv., **64** (1989), 375–394.

[81] J.N.Mather: *Action minimizing invariant measures for positive definite Lagrangian systems.* Math. Z., **207** (1991), 169–207.

[82] E.Y.Panov: *On strong precompactness of bounded sets of measure-valued solutions of a first order quasilinear equation.* Math. Sb., **186** (1995), 729–740.

[83] G.Petrova & B.Popov: *Linear transport equation with discontinuous coefficients.* Comm. PDE, **24** (1999), 1849–1873.

[84] F.Poupaud & M.Rascle: *Measure solutions to the liner multidimensional transport equation with non-smooth coefficients.* Comm. PDE, **22** (1997), 337–358.

[85] A.Pratelli: *Equivalence between some definitions for the optimal transport problem and for the transport density on manifolds.* Ann. Mat. Pura Appl., **184** (2005), 215–238.

[86] S.K.Smirnov: *Decomposition of solenoidal vector charges into elementary solenoids and the structure of normal one-dimensional currents.* St. Petersburg Math. J., **5** (1994), 841–867.

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

[88] L. Tartar: *Compensated compactness and applications to partial differential equations.* Research Notes in Mathematics, Nonlinear Analysis and Mechanics, ed. R. J. Knops, vol. **4**, Pitman Press, New York, 1979, 136–211.

[89] R.Temam: *Problémes mathématiques en plasticité.* Gauthier-Villars, Paris, 1983.

[90] J.I.E.Urbas: *Global Hölder estimates for equations of Monge-Ampère type*, Invent. Math., **91** (1988), 1–29.

[91] J.I.E.Urbas: *Regularity of generalized solutions of Monge-Ampère equations*, Math. Z., **197** (1988), 365–393.

[92] A.Vasseur: *Strong traces for solutions of multidimensional scalar conservation laws.* Arch. Ration. Mech. Anal., **160** (2001), 181–193.

[93] C.Villani: *Topics in mass transportation.* Graduate Studies in Mathematics, **58** (2004), American Mathematical Society.

[94] C.Villani: *Optimal transport: old and new.* Lecture Notes of the 2005 Saint-Flour Summer school.

[95] L.C.Young: *Lectures on the calculus of variations and optimal control theory,* Saunders, 1969.