Transport equation and Cauchy problem for BV vector fields

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

1 Introduction

In this paper we extend to the case of BV vector fields B the DiPerna–Lions theory [26], showing that locally bounded weak solutions w of the PDE $B \cdot \nabla w = c$ in an open set A in $\mathbb{R}^{d+1}$ are renormalizable whenever $B \in BV_{\rm loc}$ and its distributional divergence belongs to $L^1_{\rm loc}$. In the particular case when $A = (0,T) \times \mathbb{R}^d$ and $B(t,x) = (1, b(t,x))$, due to the special structure of the vector field B we need only BV regularity of b with respect to the spatial variables and we obtain uniqueness and comparison results for bounded and compactly supported (in space) solutions of the transport equation

$$\frac{\partial w}{\partial t}(t,x) + b(t,x) \cdot \nabla w(t,x) = e(t,x) w(t,x) \qquad (t,x) \in (0,T) \times \mathbb{R}^d \tag{1.1}$$

when $b(t,\cdot) \in BV_{\rm loc}(\mathbb{R}^d)$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$ and

$$\int_0^T \|e(t,\cdot)\|_{L^\infty(B_R)} + \|{\rm div}\, b(t,\cdot)\|_{L^\infty(B_R)} + \|b(t,\cdot)\|_{BV(B_R)}\, dt < +\infty \tag{1.2}$$

for any $R > 0$ (here ${\rm div}\, b$ refers to the spatial variables only and $\mathscr{L}^m$ denotes the Lebesgue measure in $\mathbb{R}^m$). Finally, we use these results to analyze to what extent the ODE $\dot\psi = b(t,\psi)$ has a unique solution for $\mathscr{L}^d$-almost every initial condition x. In particular, when b is also globally bounded we define a class of regular Lagrangian flows, proving for them existence, uniqueness, stability and injectivity results.

As in [26], the key ingredient leading to the stability of the PDE $B \cdot \nabla w = c$ (here $B \cdot \nabla w$ is understood as a distribution, see (3.3)) under renormalization is a smoothing argument and the analysis of the commutator

$$r_\delta := (B \cdot \nabla w) * \rho_\delta - B \cdot \nabla (w * \rho_\delta). \tag{1.3}$$

It was shown in [26] that $r_\delta \to 0$ in $L^1_{\rm loc}$ as $\delta \downarrow 0$ when $w \in L^\infty_{\rm loc}$ and $B \in W^{1,1}_{\rm loc}$. It turns out that the asymptotic behaviour of the commutator is very sensitive to the choice of the convolution kernel ρ: for instance it was noticed in [14] that if ρ is a radial convolution kernel, then still $r_\delta \to 0$ in $L^1_{\rm loc}$ provided

$$D_i B^j + D_j B^i \in L^1_{\rm loc} \qquad \forall i,\, j = 1, \ldots, d. \tag{1.4}$$

This is quite surprising, as no information on the antisymmetric part of the distributional derivative is needed. However condition (1.4), unlike the BV or Sobolev property, is not invariant under bi-Lipschitz changes of coordinates and it is not fulfilled in several cases of interest (see for instance the Vlasov equation considered by F.Bouchut in [12], or the canonical I order Hamiltonian reformulation of II order ODE's) unless $B \in W^{1,1}$.

A further progress on this problem was achieved in the papers [31], [12], [19]. In [31] P.L.Lions extends the results of [26] to a suitable class of "piecewise $W^{1,1}$ functions". In [12] F.Bouchut extends the theory to equations of Vlasov type using the special structure of the equation and the different roles played by the time variable t, the space variable x and the velocity variable ξ (see also [29] for an extension to velocity fields with point singularities). The key idea in [12], already present in a more implicit form [31], is an anisotropic smoothing argument which takes into account the particular structure of the vector field and leads to commutators with a very small L^1 norm. This idea is further developed by F.Colombini and N.Lerner in [19], where they define a class of "conormal BV vector fields", invariant under bi-Lipschitz transformations and including the class considered in [31] (see also [18] for uniqueness results for continuous solutions) and prove a uniqueness result for bounded solutions, assuming that the boundary condition is prescribed on a transversal hypersurface (in a suitable sense) to B. The main assumption in [19] is that, possibly excluding a relatively closed and $\mathscr{H}^{d-1}$-negligible subdomain, locally the derivative of the vector field is a measure only along one direction, and there is absolute continuity along the other ones.

Our starting point has been the intuition that, roughly speaking, *any* BV vector field u asymptotically behaves on a small scale as one of the vector fields considered in [19]: this is a consequence of Alberti's rank one theorem [2], saying that the singular part $D^s u$ with respect to $\mathscr{L}^d$ of the distributional derivative Du of a function $u \in BV_{\rm loc}(\mathbb{R}^d;\mathbb{R}^m)$ can be written, in polar decomposition, as $D^s u = \xi \otimes \eta |D^s u|$. By a standard blow-up argument (see for instance Theorem 3.95 in [6]) this implies that near to $|D^s u|$-a.e. point x asymptotically $u(y)$ behaves like a function having a single non-zero component, parallel to $\eta(x) \in \mathbf{S}^{m-1}$, and depending on a single scalar variable, the component of y along $\xi(x) \in \mathbf{S}^{d-1}$. Moreover, when $m = d$ the assumption that the distributional divergence of u is absolutely continuous with respect to $\mathscr{L}^d$ implies that $\xi(x) \perp \eta(x)$ for $|D^s u|$-a.e. x.

Assuming for simplicity $\|w\|_\infty \leq 1$ and denoting by $D^a u$ the absolutely continuous part of the derivative, we prove that for any compact set $Q \subset \Omega$ the quantity

$\limsup_{\delta\downarrow 0}\int_Q |r_\delta|\,dx$ can be estimated either by

$$\int_Q\int_{\mathbb{R}^d} |\langle\xi(x),z\rangle||\langle\eta(x),\nabla\rho(z)\rangle|\,dz\,d|D^s u|(x)+\left(d+\int_{\mathbb{R}^d}|z||\nabla\rho(z)|\,dz\right)|D^a u|(Q)$$

or by

$$\left(\int_{\mathbb{R}^d}|z||\nabla\rho(z)|\,dz\right)|D^s u|(Q).$$

Roughly speaking, the first estimate is useful in the regions where the singular part of the distributional derivative is dominant, while the second one is useful in the regions where the absolutely continuous part is dominant. Moreover, using the information that $\xi\perp\eta$, for $|D^s u|$-a.e. x we can properly choose the convolution kernel ρ in such a way that (see Lemma 3.3)

$$\int_{\mathbb{R}^d}|\langle\xi(x),z\rangle||\langle\eta(x),\nabla\rho(z)\rangle|\,dz << 1.$$

We are able in §3 to combine these estimates for the functions $u=b(t,\cdot)$ into a suitable localization and regularization scheme (only with respect to the spatial variables) to obtain the renormalization property. We would like also to stress that the optimal regularity class for b ensuring the validity of the renormalization property is not known; the examples in [25], [20], [1], [20] indicate that in general the BV condition could be close to be sharp. On the other hand, there are examples of non BV vector fields for which the renormalization property is true, for instance those of the form $B=(\rho,\rho b)$, with $D\cdot B=0$ and $\rho>0$, $\rho+1/\rho\in L^\infty$, $b\in BV_{\rm loc}$ and $D_x\cdot b\ll\mathscr{L}^d$ (see [9], notice that no BV regularity on ρ is imposed).

In §4, following a by now standard argument, we obtain some uniqueness and comparison results for bounded solutions of (1.1). To keep the treatment of this topic as much simple as possible we consider only solutions with compact support in space. This restriction can be weakened in many ways, along the duality theory developed in [26] (see also [9]), but compactly supported functions suffice to analyze the problem from the Lagrangian viewpoint, one of the final goals of this paper.

While the first part of this paper is a technical improvement of [26], in the second part, concerned with the analysis of uniqueness of solutions from the Lagrangian viewpoint, we depart from the methods of [26] and we introduce new tools, borrowed from Probability and Calculus of Variations.

In §5 we consider solutions of the transport equation induced by characteristics, in a measure-theoretic framework. Specifically, given a $\mathscr{L}^d$-measurable family of measures γ_x in $C([0,T];\mathbb{R}^d)$, we show that the measures

$$\mu_t^{\gamma_x}(\varphi):=\int_{\mathbb{R}^d}\int_{C([0,T];\mathbb{R}^d)}\varphi(\psi)\,d\gamma_x(\psi)\,dx\qquad\qquad \varphi\in C_c(\mathbb{R}^d)\tag{1.5}$$

solve the transport equation (1.1) with $e = -\mathrm{div}\, b(t,\cdot)$ and with the initial condition $\mu_0 = \gamma_x(C([0,T];\mathbb{R}^d))\mathscr{L}^d$ (the equation and the initial condition make sense for measures as well) provided for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$ the measure γ_x is concentrated on the family of solutions of the ODE $\dot\psi = b(t,\psi)$ with $\psi(0) = x$. This formula reduces to the classical one for the representation of solutions of the transport equation in the case when all measures are Dirac masses, i.e. $\gamma_x = \delta_{\psi_x}$ with $\psi_x(\cdot)$ solution of the Cauchy problem relative to b. This viewpoint is very close in spirit to L.C.Young's theory of generalized solutions [34] and to [33], where solutions of the transport equation are built using, instead, generalized characteristics in the sense of Filippov. The class of solutions in (1.5) is strongly localizable, as the Lagrangian viewpoint suggests, both with respect to the initial set I and with respect to the final set F, setting

$$\gamma_x^{I,F} := \chi_I(x)\gamma_x \llcorner \{\psi :\ \psi(T) \in F\}$$

and defining a new solution $\mu_t^{I,F}$ of the transport equation as in (1.5) using $\gamma_x^{I,F}$ instead of γ_x.

Using this kind of flexibility one can obtain that the measures γ_x have to be Dirac masses for $\mathscr{L}^d$-a.e. x provided the measures $\mu_t^{B_R,\mathbb{R}^d}$ are absolutely continuous with respect to $\mathscr{L}^d$ with uniformly bounded densities. The idea of the proof is to show first a general and elementary criterion: a family of probability measures γ_x in $C([0,T];\mathbb{R}^d)$ is made by Dirac masses for $\mathscr{L}^d$-a.e. x if and only if

$$\gamma_x\left(\{\psi :\ \psi(t) \in F\}\right)\gamma_x\left(\{\psi :\ \psi(t) \in F'\}\right) = 0 \qquad \text{for } \mathscr{L}^d\text{-a.e. } x \tag{1.6}$$

for any $t \in [0,T]$ and any pair of disjoint Borel sets F, $F' \subset \mathbb{R}^d$. When γ_x is concentrated on the family of solutions of the ODE starting at x, the orthogonality property (1.6) is shown using the comparison results of §4, localizing with respect to F and using in particular the finiteness of

$$\int_0^T \|[\mathrm{div}\, \mathrm{b}_t]^-\|_{L^\infty(B_R)}\, dt \qquad \forall R > 0.$$

Under the assumption

$$\int_0^T \|[\mathrm{div}\, \mathrm{b}_t]^+\|_{L^\infty(B_R)}\, dt \qquad \forall R > 0$$

one can show, instead, that the flow $\mu_t^{I,\mathbb{R}^d}$ is injective in the following weak sense: if $I \cap I' = \emptyset$ then $\mu_t^{I,\mathbb{R}^d}$ is orthogonal to $\mu_t^{I',\mathbb{R}^d}$ for any $t \in [0,T]$ (see also Theorem 5.7 for a more detailed statement).

In §6 we study the "generic uniqueness" of the ODE $\dot\psi = b(t,\psi)$ (in this connection, see [16] and [17], where the authors deal with the ODE $\dot\psi \in A(\psi)$ with A maximal monotone).

Assuming, in addition to (1.2), that b is globally bounded, we show that for any $R > 0$ there is essentially a unique way to select for any $x \in B_R$ a solution $\psi_x(t)$ of the ODE in such a way that the measures

$$\mu_t(\varphi) := \int_{B_R} \varphi(\psi(x)(t))\, dx \qquad \varphi \in C_c(\mathbb{R}^d),\ t \in [0,T]$$

are absolutely continuous and have uniformly bounded densities. The proof uses the theory of Young measures and the regsults of the previous section. We call this selection *regular Lagrangian flow* and show that regular Lagrangian flows enjoy a natural stability property with respect to the approximation of b by vector fields b_h which are Lipschitz with respect to the x variable. This stability property is the natural extension to the BV case of the results in Section III of [26] (see also Remark 6.7 for a more precise comparison of the two approaches).

Notice however that our methods do not seem to apply for proving the *stronger* statement that for $\mathscr{L}^d$-a.e. x there is a unique solution of the ODE $\dot{\psi} = b(t,\psi)$ with $\psi(0) = x$ (as in the cases considered in [16], [17]): indeed, there is no rigorous justification of why *any* flow μ_t built by making a (measurable) selection of solutions of the ODE is made by absolutely continuous measures with uniformly bounded densities. This problem, which to the author's knowledge is open even when b is in a Sobolev space and is independent of t, seems to be related to the problem of finding a suitable approximation scheme for proving the following *formal* estimate (see Remark 6.3) for positive measures μ_t:

$$\frac{\partial \mu_t}{\partial t} + D_x \cdot (b\mu_t) = 0,\ \ \mu_0 \leq \mathscr{L}^d \quad \Longrightarrow \quad \mu_t \leq e^{\int_0^t \|\mathrm{div}\, b_s\|_\infty\, ds} \mathscr{L}^d.$$

We remark also that the validity of this stronger statement, due to the presence of the product $b\mu_t$, could be very sensitive to the choice of a good representative in the equivalence class of b (for instance the approximate limit of b, wherever this exists), while the class of regular Lagrangian flows is actually independent on the choice of this representative.

Finally, we mention some recent applications of this paper. In [8], following a suggestion by A.Bressan [13], we obtain an existence result for the system of conservation laws

$$\frac{\partial u}{\partial t} + \sum_{i=1}^d \frac{\partial}{\partial x_i} \left(f(|u|)u \right) = 0, \qquad u(0,\cdot) = \bar{u}$$

with $f : \mathbb{R} \to \mathbb{R}^m$ smooth, assuming only that $|\bar{u}| \in L^\infty \cap BV_{\rm loc}(\mathbb{R}^d)$ and $|\bar{u}| \geq c > 0$. This is the first general existence result for systems of this type when d, $m \geq 3$. Existence is shown in the class of functions u such that $\rho := |u|$ is an entropy solution of

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^d \frac{\partial}{\partial x_i} \left(f(\rho)\rho \right) = 0, \qquad u(0,\cdot) = |\bar{u}| \tag{1.7}$$

and it is obtained by the method of characteristics, i.e. solving the ODE $\dot{y} = f(\rho(t,y))$ and using Theorem 6.6. In a later paper [9], using instead the Eulerian part of this paper and some arguments from Di Perna–Lions theory, we obtain stronger results, removing the assumption $|\bar{u}| \geq c > 0$ and showing uniqueness of u in the class of functions whose modulus fulfils (1.7). We discuss also some conjectures related to stronger forms of uniqueness, and their relation with a compactness conjecture for Lagrangian flows made by A.Bressan in [13] (Theorem 6.6 provides only a partial answer to that conjecture).

More recently, still using Theorem 6.6, M.Cullen and M.Feldman were able to show in [22] existence for the semigeostrophic system arising in metereology (here $D_t = \partial_t + u \cdot \nabla$ is the Eulerian derivative)

$$\begin{cases} D_t(v_1^0, v_2^0) + (\partial_1 p, \partial_2 p) = (u_2, -u_1), \quad (v_1^0, v_2^0) = (-\partial_2 p, \partial_1 p) \\ D_t\rho = 0, \quad D \cdot u = 0, \quad \partial_3 p + \rho = 0 \end{cases}$$

in the original physical variables, using previous existence results (see [10], [21]) in the so-called dual coordinates. Since the passage from physical to dual coordinates requires a non smooth but BV vector field (more precisely, the gradient of a convex function), the stability theorem provides a natural way to justify some formal calculations.

Acknowledgement. This work is part of the European Research Training Network "Homogenization and Multiple Scales" (HMS 2000), under contract HPRN-2000-00109. It has been also supported by COFIN02-MIUR, National Project "Calculus of Variations".

2 Notation and preliminary results on BV functions

In this section we introduce our main notation and recall some basic facts about BV functions. General references on this topic are [6] and [27].

If X is a complete and separable metric space we denote by $\mathcal{B}(X)$ the σ-algebra of Borel subsets of X, by $\mathscr{M}_1(X)$ the collection of all positive measures $\mu : \mathcal{B}(X) \to [0, +\infty)$ in X with $\mu(X) \leq 1$ and by $\mathscr{P}(X)$ the subset made by probability measures. We canonically extend any positive and locally finite measure $\mu : \mathcal{B}(X) \to [0, +\infty]$ to the σ-algebra of μ-measurable sets.

If μ is a positive locally finite measure in X and $C \subset X$ is μ-measurable we define $\mu \llcorner C$ by

$$\mu \llcorner C(A) := \mu(A \cap C) \qquad \forall A \in \mathcal{B}(X).$$

The Lebesgue d-dimensional measure and the Hausdorff $(d-1)$-dimensional measure in $\mathbb{R}^d$ will be respectively denoted by $\mathscr{L}^d$ and $\mathscr{H}^{d-1}$. The relations of absolute continuity and perpendicularity between positive measures (possibly σ-finite) will be denoted respectively by $\ll$ and $\perp$.

In the following we denote by Ω a generic open set in $\mathbb{R}^d$.

Definition 2.1 (Approximate limit and approximate discontinuity set). Let $w \in L^1_{\rm loc}(\Omega;\mathbb{R}^m)$. We say that $x \in \Omega$ is an *approximate continuity* point of w if there exists $z \in \mathbb{R}^m$ such that

$$\lim_{r\downarrow 0} \frac{1}{r^d}\int_{B_r(x)} |w(y)-z|\,dy = 0.$$

The vector z above, called *approximate limit* of w at x, is unique and it will be denoted by $\tilde{w}(x)$. We denote by $S_w \in \mathcal{B}(\mathbb{R}^d)$ the complement of the approximate continuity set of w.

Notice that by Lebesgue differentiation theorem the set S_w is $\mathscr{L}^d$-negligible and $\tilde{w} = w$ $\mathscr{L}^d$-a.e. in $\Omega \setminus S_w$. Recall also that the mollified functions $w_\epsilon = w * \rho_\epsilon$ converge pointwise out of S_w, and the pointwise limit is $\tilde{w}$.

Let $u \in L^1_{\rm loc}(\Omega;\mathbb{R}^m)$. We denote by $Du = (D_i u^j)$ the derivative in the sense of distributions of u, the $\mathbb{R}^{m\times d}$-valued distribution defined by

$$D_i u^j(\varphi) := -\int_\Omega u^j \frac{\partial \varphi}{\partial x_i}\,dx \quad \forall \varphi \in C^\infty_c(\Omega), \qquad 1 \le i \le d,\ 1 \le j \le m.$$

Definition 2.2 (BV functions). We say that $u \in L^1(\Omega;\mathbb{R}^m)$ has *bounded variation* in Ω is Du is representable by a $\mathbb{R}^{m\times d}$-valued measure, still denoted with Du, with finite total variation in Ω.

We recall that the *total variation* $|\lambda|$ of a $\mathbb{R}^{d\times m}$-valued measure λ is defined by

$$|\lambda|(C) := \sup\left\{\sum_{i=1}^\infty |\lambda(C_i)| :\ C_i \in \mathcal{B}(\Omega) \text{ pairwise disjoint, } C_i \subset C\right\}$$

(choosing for convenience the Hilbert–Schmidt norm in $\mathbb{R}^{m\times d}$) and that if $|\lambda|(\Omega)$ is finite then λ admits the *polar decomposition* $\lambda = M|\lambda|$ with $|M(x)| = 1$ for $|\lambda|$-a.e. $x \in \Omega$.

We consider the canonical decomposition of Du as $D^a u + D^s u$, with $|D^a u| \ll \mathscr{L}^d$ and $|D^s u| \perp \mathscr{L}^d$ and we denote by $\nabla u = (\partial u^j/\partial x_i) \in \mathbb{R}^{m\times d}$ the density of $D^a u$ with respect to $\mathscr{L}^d$ (if $d = 1$ we use also the notation u').

The following deep result (see [2] and also [3] for a proof that works for gradients of convex functions only) is one of the main tools used in this paper (see also Remark 3.7).

Theorem 2.3 (Alberti's rank one theorem). *Let $u \in BV(\Omega;\mathbb{R}^m)$ and let $D^s u = M|D^s u|$ be the polar decomposition of the singular part of its distributional derivative. Then $M(x)$ has rank one, i.e. $M(x) = \eta(x) \otimes \xi(x)$, with $|\xi(x)| = |\eta(x)| = 1$, for $|D^s u|$-a.e. $x \in \Omega$.*

Recall that for $\xi \in \mathbb{R}^d$ and $\eta \in \mathbb{R}^m$ the product $\eta \otimes \xi$ is the linear map from $\mathbb{R}^d$ to $\mathbb{R}^m$ defined by $\eta \otimes \xi(v) = \langle \xi, v\rangle \eta$. Notice that the pair of unit vectors $\xi(x)$, $\eta(x)$ is uniquely determined $|D^s u|$-a.e. up to a change of sign. On the jump set of u the vector $\xi(x)$ is the (approximate) normal to the jump set and $\eta(x)$ is the direction of the jump. This fact can be proved by an elementary blow-up argument, but the non-trivial content of Theorem 2.3 is that the same rank-one structure is retained also by the diffuse part of $D^s u$, the so-called Cantor part of derivative (see Chapter 3 in [6]).

In the particular case $m = d$ we can write the *distributional divergence* $D \cdot u$ as

$$D \cdot u = \mathrm{trace}(\nabla u)\mathscr{L}^d + \langle \xi, \eta \rangle |D^s u|,$$

hence

$$D \cdot u \ll \mathscr{L}^d \qquad \text{if and only if} \qquad \xi \perp \eta \quad |D^s u|\text{-a.e. in } \Omega. \tag{2.1}$$

We also recall that for any vector $z \in \mathbb{R}^d$ and any compact set $K \subset \Omega$ we have

$$\int_K |u(x+z) - u(x)| \, dx \le |\sum_{i=1}^d z_i D_i u|(K_{|z|}), \tag{2.2}$$

where $K_{|z|}$ is the open $|z|$-neighbourhood of K, provided $|z| < \mathrm{dist}(K, \partial\Omega)$. The following theorem provides more information on the behaviour of the increment $u(x+z) - u(x)$ as $|z| \to 0$.

Theorem 2.4 (Difference quotients of BV functions). *Let $u \in BV_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^m)$ and let $z \in \mathbb{R}^d$. Then the difference quotients*

$$\frac{u(x+\delta z) - u(x)}{\delta}$$

can be canonically written as $u^1_\delta(z) + u^2_\delta(z)$, where $u^1_\delta(z)$ strongly converge as $\delta \downarrow 0$ in $L^1_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^m)$ to

$$\sum_{i=1}^d \frac{\partial u}{\partial x_i}(x) z_i$$

and

$$\limsup_{\delta \downarrow 0} \int_K |u^2_\delta(z)| \, dx \le |z| |D^s u|(K) \tag{2.3}$$

for any compact set $K \subset \mathbb{R}^d$. In addition we have the uniform bound

$$\sup_{z \in K'} \sup_{\delta \in (0,\epsilon)} \int_K |u^1_\delta(z)| + |u^2_\delta(z)| \, dx \le \sup_{z \in K'} |z| |Du| \left(\{x : \ \mathrm{dist}(x, K) \le \epsilon\}\right) \tag{2.4}$$

whenever K, K' are compact subset of $\mathbb{R}^d$ and $\epsilon > 0$.

Proof. Setting $u^i_\delta(z) = u^i_{\delta|z|}(z/|z|)$ we can reduce ourselves to the case when z is a unit vector. Up to a rotation we assume that $z = e_d$ and we write $x = (x', x_d)$ with $x' \in \mathbb{R}^{d-1}$ and $x_d \in \mathbb{R}$.

We first recall some elementary facts from Real Analysis: if μ is a $\mathbb{R}^m$-valued measure in $\mathbb{R}$ with locally finite variation, then by Fubini's theorem the functions

$$\hat{\mu}_\delta(t) := \frac{\mu([t, t+\delta])}{\delta} = \mu * \frac{\chi_{[0,\delta]}}{\delta}(t), \qquad t \in \mathbb{R}$$

satisfy

$$\int_K |\hat{\mu}_\delta| \, dt \le |\mu|(K_\delta) \qquad \text{for any compact set } K \subset \mathbb{R}, \tag{2.5}$$

where K_δ is the open δ neighbourhood of K. A density argument based on (2.5) then shows that $\hat{\mu}_\delta$ converge in $L^1_{\rm loc}(\mathbb{R})$ to the density of μ with respect to $\mathscr{L}^1$ whenever $\mu \ll \mathscr{L}^1$.

We define

$$u^1_\delta(x', x_d) := \widehat{\frac{\partial u}{\partial x_d}(x', \cdot)\mathscr{L}^1}(x_d) = \frac{1}{\delta}\int_{x_d}^{x_d+\delta} \frac{\partial u}{\partial x_d}(x', s)\, ds.$$

Fubini's theorem and (2.5) imply the $L^1_{\rm loc}(\mathbb{R}^d; \mathbb{R}^m)$ convergence of u^1_δ to $\partial u/\partial x_d$.

It remains to show the estimate (2.3). Setting $u_{x'}(s) := u(x', s)$, we recall that (see for instance Theorem 3.103, Theorem 3.107 and (3.108) in [6]) for $\mathscr{L}^{d-1}$-a.e. x' the following properties hold:

(a) $u_{x'} \in BV_{\rm loc}(\mathbb{R}; \mathbb{R}^m)$;

(b) $u'_{x'}(s) = \partial u/\partial x_d(x', s)$ for $\mathscr{L}^1$-a.e. $s \in \mathbb{R}$;

(c) for any $\delta > 0$ we have $u_{x'}(s+\delta) - u_{x'}(s) = Du_{x'}([s, s+\delta])$ for $\mathscr{L}^1$-a.e. $s \in \mathbb{R}$.

Moreover, we have also $\int_{\mathbb{R}^{d-1}} |D^s u_{x'}|\, dx' \le |D^s u|$. Therefore, given any $\delta > 0$, for $\mathscr{L}^{d-1}$-a.e. x' we have

$$\begin{aligned}
\frac{u(x', x_d+\delta) - u(x', x_d)}{\delta} &= \frac{u_{x'}(x_d+\delta) - u_{x'}(x_d)}{\delta} = Du_{x'}([x_d, x_d+\delta]) \\
&= \widehat{u'_{x'}\mathscr{L}^1}(x_d) + \widehat{D^s u_{x'}}(x_d) \\
&= u^1_\delta(x', x_d) + \widehat{D^s u_{x'}}(x_d) \quad \text{for } \mathscr{L}^1\text{-a.e. } x_d.
\end{aligned}$$

By (2.5) it follows that

$$\begin{aligned}
\int_K |u^2_\delta|\, dx_d\, dx' &\le \int_{\mathbb{R}^{d-1}} \int_{\{x_d:\ (x', x_d) \in K\}} |\widehat{D^s u_{x'}}|(x_d)\, dx_d\, dx' \\
&\le \int_{\mathbb{R}^{d-1}} |D^s u_{x'}|(\{x_d :\ (x', x_d) \in K_\delta\})\, dx' \le |D^s u|(K_\delta).
\end{aligned}$$

□

3 Stability of the transport equation under renormalization

In this section we fix $T \in (0,+\infty)$, an open set $\Omega \subset \mathbb{R}^d$ and consider a function $b(t,x) : \Omega \times (0,T) \to \mathbb{R}^d$ satisfying the following conditions:

$$b_t(x) := b(t,x) \in BV_{\mathrm{loc}}(\Omega) \text{ for } \mathscr{L}^1\text{-a.e. } t \in (0,T); \tag{3.1}$$

$$\int_{I\times B_R} |b_t|\,dx\,dt + \int_I |Db_t|(B_R)\,dt < +\infty \qquad \text{whenever } R>0 \text{ and } \overline{I} \subset (0,T). \tag{3.2}$$

We assume also that for $\mathscr{L}^1$-a.e. $t \in (0,T)$ the distribution $D\cdot b_t$ is representable as $\mathrm{div}\, b_t \mathscr{L}^d$.

In the following theorem and in the sequel we use the fact that the distribution $b\cdot\nabla w$ is well defined in $(0,T)\times\Omega$ when $w \in L^\infty_{\mathrm{loc}}((0,T)\times\Omega)$, by the identity

$$\begin{aligned}\langle b\cdot\nabla w, \varphi\rangle &:= \langle D_x\cdot(wb),\varphi\rangle - \langle wD_x\cdot b,\varphi\rangle \\ &= -\int_{(0,T)\times\mathbb{R}^d} w\nabla\varphi\cdot b\,dt\,dx - \int_{(0,T)\times\mathbb{R}^d} w\varphi \mathrm{div}\, b_t\,dt\,dx\end{aligned} \tag{3.3}$$

for any $\varphi \in C^\infty_c((0,T)\times\Omega)$.

We write

$$D^s b_t = M_t |D^s b_t| \qquad \text{for } \mathscr{L}^1\text{-a.e. } t\in(0,T) \tag{3.4}$$

and apply Alberti's theorem and (2.1) to obtain vectors $\eta_t(x)$, $\xi_t(x) \in \mathbf{S}^{d-1}$ such that

$$M_t = \eta_t \otimes \xi_t \quad \text{and} \quad \xi_t \perp \eta_t \qquad |D^s b_t|\text{-a.e.} \tag{3.5}$$

for $\mathscr{L}^1$-a.e. $t\in(0,T)$. By integration in time we define $|Db| = \int_0^T |Db_t|\,dt$, so that

$$\int_{(0,T)\times\Omega} \varphi(t,x)\,|Db|(t,x) = \int_0^T \int_\Omega \varphi(t,x)\,d|Db_t|(x)\,dt \qquad \forall\varphi\in C_c((0,T)\times\Omega).$$

We apply a similar construction to define $|D^a b|$, absolutely continuous with respect to $\mathscr{L}^{d+1}$, and $|D^s b|$, singular with respect to $\mathscr{L}^{d+1}$.

Definition 3.1 (Anisotropic and isotropic energy of a kernel). *For any $d\times d$ matrix M and any convolution kernel ρ with compact support we define*

$$\Lambda(M,\rho) := \int_{\mathbb{R}^d} |\langle Mz, \nabla\rho(z)\rangle|\,dz.$$

We define also

$$I(\rho) := \int_{\mathbb{R}^d} |z||\nabla\rho(z)|\,dz.$$

The following theorem provides two quantitative estimates of the commutator appearing in Lemma II.1 of [26] in the case when the vector field has bounded variation with respect to the spatial variables. Notice that the mollification applies to these variables only.

Theorem 3.2 (Commutator estimate). *Assume (3.1), (3.2) and that $D \cdot b_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$. For any $w \in L^\infty_{\rm loc}((0,T)\times\Omega)$ and for any even convolution kernel $\rho : \mathbb{R}^d \to [0,+\infty)$ with $\operatorname{supp}\rho \subset B_1$ define*

$$r_\delta := (b \cdot \nabla w) * \rho_\delta - b \cdot \nabla (w * \rho_\delta).$$

Let $Q \subset A \subset\subset (0,T)\times\Omega$, with Q compact and A open, and let $L := \|w\|_{L^\infty(A)}$. Then

$$\limsup_{\delta\downarrow 0} \int_Q |r_\delta|\,dt\,dx \le L \int_Q \Lambda(M_t(x),\rho)\,d|D^s b|(t,x) + L\,(d + I(\rho))\,|D^a b|(Q) \tag{3.6}$$

(with M_t defined in (3.4)) and

$$\limsup_{\delta\downarrow 0} \int_Q |r_\delta|\,dt\,dx \le L I(\rho)|D^s b|(Q). \tag{3.7}$$

Proof. Let $\tau = \operatorname{dist}(Q,\partial A)$ and assume with no loss of generality that $\|w\|_{L^\infty(A)} = 1$. We use first the eveness of the kernel ρ to derive the identity

$$r_\delta(t,x) = \int_{\mathbb{R}^d} w(t,y)\,\{(b_t(y) - b_t(x)) \cdot \nabla\rho_\delta(x-y)\}\,dy - (w\operatorname{div} b_t) * \rho_\delta(x) \tag{3.8}$$

for $(t,x) \in Q$ and $\delta < \tau$. Indeed, testing the right hand side of (3.8) with $\varphi \in C^\infty_c(A)$ satisfying $\operatorname{dist}(\operatorname{supp}\varphi, \partial A) \ge \delta$ and using (3.3) we obtain

$$\begin{aligned} & -\langle wb, \nabla(\varphi * \rho_\delta)\rangle - \langle b \cdot \nabla(w * \rho_\delta), \varphi\rangle - \langle w\operatorname{div} b_t, \varphi * \rho_\delta\rangle \\ = \; & \langle b \cdot \nabla w, \varphi * \rho_\delta\rangle - \langle b \cdot \nabla(w * \rho_\delta), \varphi\rangle = \langle r_\delta, \varphi\rangle. \end{aligned}$$

For $\delta < \tau$ and $(t,x) \in Q$ we change variables in (3.8) to obtain

$$r_\delta(t,x) = \int_{\mathbb{R}^d} w(t, x-\delta z)\left\{\frac{b_t(x-\delta z) - b_t(x)}{\delta} \cdot \nabla\rho(z)\right\}dz - (w\operatorname{div} b_t) * \rho_\delta(t,x).$$

The proof of (3.7) follows from applying Theorem 2.4 to the functions b_t and splitting as in the statement of the theorem the difference quotient in two parts:

$$\begin{aligned} & \lim_{\delta\downarrow 0} \int_Q \left| \int_{\mathbb{R}^d} w(t,x-\delta z)\left\{\frac{b_t(x-\delta z) - b_t(x)}{\delta} \cdot \nabla\rho(z)\right\}dz - (w\operatorname{div} b_t) * \rho_\delta(t,x)\right| dt\,dx \\ \le \; & \lim_{\delta\downarrow 0} \int_Q \left| \int_{\mathbb{R}^d} w(t,x-\delta z)(b_t)^1_\delta(-z) \cdot \nabla\rho(z)\,dz - w(t,x)\operatorname{div} b_t(t,x)\right| dt\,dx \\ + \; & \limsup_{\delta\downarrow 0} \int_{\mathbb{R}^d} |\nabla\rho(z)| \int_Q |(b_t)^2_\delta(-z)|\,dx\,dt\,dz. \end{aligned}$$

The first limit is equal to 0 because

$$w(t,x-\delta z)(b_t)^1_\delta(-z)(x) \to w(t,x)\nabla b_t(x)(-z) \qquad \text{strongly in } L^1_{\rm loc}(\Omega\times\mathbb{R}^d)$$

for $\mathscr{L}^1$-a.e. $t\in(0,T)$ (taking into account the strong $L^1_{\rm loc}(\mathbb{R}^d)$ convergence of $(b_t)^1_\delta(-z)$, (2.4) and the strong continuity in any $L^p_{\rm loc}$, $p<+\infty$, of translations), so that the limit equals:

$$\int_Q \left| w(t,x)\left(\int_{\mathbb{R}^d}\sum_{i,j=1}^d \frac{\partial b_t^j}{\partial x_i}(x) z_i \frac{\partial\rho}{\partial z_j}(z)\,dz + \operatorname{div} b_t(x)\right)\right|\,dt\,dx = 0$$

using the fact that $\int z_i\partial\rho/\partial z_j\,dz = -\delta_{ij}$. The lim sup can be estimated with $I(\rho)|D^s b|(Q)$ using (2.3) and (2.4).

Now we show (3.6). In this case we do not take advantage of the cancellations in the z variable and estimate:

$$\|r_\delta\|_{L^1(Q)} \le \int_{\mathbb{R}^d}\int_Q \left|\frac{b_t(x-\delta z)-b_t(x)}{\delta}\cdot\nabla\rho(z)\right|\,dt\,dx\,dz + \|\operatorname{div} b_t\|_{L^1(Q_\delta)}, \tag{3.9}$$

where Q_δ is the open δ-neighbourhood of Q. Let us consider for any $z\in\mathbb{R}^d$ fixed the functions $v_{tz} = b_t\cdot\nabla\rho(z)$, whose distributional derivative is given by

$$D_i v_{tz} = \sum_{j=1}^d \frac{\partial b_t^j}{\partial x_i}\frac{\partial\rho}{\partial z_j}(z)\mathscr{L}^d + (M_t)_{ji}\frac{\partial\rho}{\partial z_j}(z)|D^s b_t|$$

for any $i=1,\ldots,d$. By applying (2.2) to v_{tz} we get:

$$\limsup_{\delta\downarrow 0}\int_Q\left|\frac{b_t(x-\delta z)-b_t(x)}{\delta}\cdot\nabla\rho(z)\right|\,dt\,dx \le \int_0^T |\sum_i z_i D_i v_{tz}|(\{x:\,(t,x)\in Q\})\,dt$$

$$\le \int_Q |\nabla b_t||z||\nabla\rho(z)|\,dt\,dx + \int_Q |\langle M_t(x)(z),\nabla\rho(z)\rangle|\,d|D^s b|(t,x).$$

Then, Fatou's Lemma and (3.9) give (3.6), using the inequality $|\operatorname{div} b_t|\mathscr{L}^d \le d|D^a b_t|$. □

As in [12] and in [19] we can try to get a small norm of the commutator by optimizing the choice of the convolution kernel. Looking at (3.6), we are led to the following problem:

$$\text{minimize } \rho\in C^\infty_c(\mathbb{R}^d)\mapsto \Lambda(M,\rho) \text{ subject to the constraints } \rho\ge 0,\ \int\rho = 1. \tag{3.10}$$

Notice that, by scaling invariance, the problem is unchanged if the additional constraint $\operatorname{supp}\rho\subset B_1$ is added. The following Lemma shows that this minimization problem has a simple solution when $M=\eta\otimes\xi$ with $\eta\perp\xi$.

Lemma 3.3 (Anisotropic convolution kernels). *Given $\xi, \eta \in \mathbf{S}^{d-1}$ with $\xi \perp \eta$ and given any $\epsilon > 0$ we can find a smooth, even convolution kernel ρ with compact support such that*

$$\int_{\mathbb{R}^d} |\langle z, \xi\rangle||\langle \nabla\rho(z), \eta\rangle|\, dz < \epsilon.$$

In particular $\Lambda(\eta\otimes\xi, \rho) < \epsilon$.

Proof. If $d = 2$ we can assume, up to a rotation, that (ξ, η) is the canonical basis of $\mathbb{R}^2$. We consider the rectangle $R_\epsilon := [-\epsilon/2, \epsilon/2] \times [-1/2, 1/2]$ and the kernel $\rho_\epsilon := \chi_{R_\epsilon}/\epsilon$. Let h_δ be the mollifiers induced by any convolution kernel h; using the fact that

$$\left|\frac{\partial(\rho_\epsilon * h_\delta)}{\partial z_i}\right| \to \frac{|\nu_i|}{\epsilon}\mathscr{H}^1 \llcorner \partial R_\epsilon \quad \text{as } \delta\downarrow 0, \qquad i = 1, 2$$

in the sense of distributions (here $\nu = (\nu_1, \nu_2)$ is the inner unit normal to R_ϵ), a straightforward computation gives

$$\lim_{\delta\downarrow 0}\int_{\mathbb{R}^2} |z_1| \left|\frac{\partial(\rho_\epsilon * h_\delta)}{\partial z_2}\right| dz = \frac{2}{\epsilon}\int_{-\epsilon/2}^{\epsilon/2} |z_1|\, dz_1 = \frac{\epsilon}{2}.$$

Therefore we can choose as ρ the function $\rho_\epsilon * h_\delta$ for $\delta = \delta(\epsilon)$ small enough.

If $d > 2$ we simply multiply the 2-dimensional kernel by a fixed kernel depending on $(d-2)$ variables orthogonal to ξ and η and compactly supported. □

We can now prove that weak solutions of the transport equation are *renormalizable*, according to the terminology introduced in [26]. In view of the applications described in [9], we extend this concept also to systems of PDE's.

Definition 3.4 (Renormalized solutions). *Assume $D\cdot b_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in (0, T)$ and that $|b_t| + |\mathrm{div}\, b_t| \in L^1_{\mathrm{loc}}((0, T) \times \Omega)$. Let $w \in L^\infty_{\mathrm{loc}}((0, T) \times \Omega; \mathbb{R}^m)$ be satisfying the system*

$$\frac{\partial w_i}{\partial t} + b\cdot\nabla w_i = c_i\mathscr{L}^{d+1} \qquad \text{in } \mathcal{D}'((0, T) \times \Omega) \tag{3.11}$$

for $i = 1, \ldots, m$, for some $c = (c_1, \ldots, c_m) \in L^1_{\mathrm{loc}}((0, T) \times \Omega; \mathbb{R}^m)$. We say that w is a renormalized solution *if*

$$\frac{\partial\beta(w)}{\partial t} + b\cdot\nabla\beta(w) = \sum_{i=1}^m c_i\frac{\partial\beta}{\partial y_i}(w)\mathscr{L}^{d+1} \qquad \text{in } \mathcal{D}'((0, T) \times \Omega)$$

for any C^1 function $\beta : \mathbb{R}^m \to \mathbb{R}$.

Notice that if w is renormalizable then all components w_i are renormalizable, but a priori the converse need not be true.

The main result of this paper, stated below, shows that if the vector field b is BV with respect to the spatial variables, then any distributional solution is a renormalized solution. The idea of the proof is to perform a localization of the problem, choosing properly in any region of spacetime where the singular part of the spatial derivative of b_t lives a convolution kernel related to the two directions (in the domain and in the codomain) where the derivative is maximal. In particular we will apply (3.6) in the regions where the singular part of the derivative is dominant and (3.7) in the regions where the absolutely continuous part of the derivative is dominant.

Theorem 3.5 (Renormalization property). *Assume (3.1), (3.2) and that* $D\cdot b_t \ll \mathscr{L}^d$ *for* $\mathscr{L}^1$*-a.e.* $t \in (0,T)$. *Let* $w \in L^\infty_{\rm loc}((0,T)\times\Omega;\mathbb{R}^m)$ *be satisfying (3.11) for some* $c \in L^1_{\rm loc}((0,T)\times\Omega;\mathbb{R}^m)$. *Then* w *is a renormalized solution.*

Proof. Since the statement to be proved is of a local nature, we fix an open domain $A \subset\subset (0,T)\times\Omega$ and set $\delta_0 := \operatorname{dist}(A, \partial(0,T)\times\Omega)$. We denote also by L the L^∞ norm of w in the $\delta_0/2$ neighbourhood of A.

Step 1. In this step we show that $\partial\beta(w)/\partial t + b\cdot\nabla\beta(w)$ is a signed measure with finite total variation in A. To this aim, we fix a smooth, even convolution kernel ρ with $\operatorname{supp}\rho \subset B_1$ and define the commutator r_δ as in Theorem 3.2, arguing component by component. The function $r_\delta = (r_{\delta,1},\ldots,r_{\delta,m}) : A \to \mathbb{R}^m$ is well-defined, provided $\delta < \delta_0$.

Notice that $w*\rho_\delta(t,x)$ are smooth with respect to the x variable in A for $\delta<\delta_0$ and

$$\begin{aligned}\frac{\partial(w_i*\rho_\delta)}{\partial t} &= \frac{\partial w_i}{\partial t}*\rho_\delta = -(b\cdot\nabla w_i)*\rho_\delta + c_i*\rho_\delta\mathscr{L}^{d+1}\\ &= -r_{\delta,i}\mathscr{L}^{d+1} - b\cdot\nabla(w_i*\rho_\delta) + c_i*\rho_\delta\mathscr{L}^{d+1} \ll \mathscr{L}^{d+1} \quad \text{in } \mathcal{D}'(A)\end{aligned}$$

for $i=1,\ldots,m$, therefore $w*\rho_\delta \in W^{1,1}(A;\mathbb{R}^m)$ and the standard chain rule can be applied. We have then the following identity:

$$\begin{aligned}&\frac{\partial\beta(w*\rho_\delta)}{\partial t} + b\cdot\nabla\beta(w*\rho_\delta) \qquad (3.12)\\ =\ & \sum_{i=1}^m \frac{\partial\beta}{\partial y_i}(w*\rho_\delta)\left[\frac{\partial(w_i*\rho_\delta)}{\partial t} + b\cdot\nabla(w_i*\rho_\delta)\right]\\ =\ & \sum_{i=1}^m \frac{\partial\beta}{\partial y_i}(w*\rho_\delta)\left[\frac{\partial w_i}{\partial t}*\rho_\delta + (b\cdot\nabla w_i)*\rho_\delta - r_{\delta,i}\mathscr{L}^{d+1}\right]\\ =\ & \sum_{i=1}^m \frac{\partial\beta}{\partial y_i}(w*\rho_\delta)\left[c_i*\rho_\delta\mathscr{L}^{d+1} - r_{\delta,i}\mathscr{L}^{d+1}\right].\end{aligned}$$

Since $\nabla\beta(w*\rho_\delta)$ is uniformly bounded in $L^\infty(A)$ and $|r_\delta|$ is uniformly bounded in $L^1(A)$ as $\delta\downarrow 0$ (using any of the estimates in Theorem 3.2), passing to the limit as $\delta\downarrow 0$ we obtain that $\partial\beta(w)/\partial t + b\cdot\nabla\beta(w)$ is a signed measure with finite total variation in A.
STEP 2. In this step we show that, inside A, the measure

$$\sigma := \frac{\partial\beta(w)}{\partial t} + b\cdot\nabla\beta(w) - \sum_{i=1}^{d}\frac{\partial\beta}{\partial y_i}(w)c_i\mathscr{L}^{d+1}$$

is singular with respect to $\mathscr{L}^{d+1}$. Notice first that the functions $\partial\beta/\partial y_i(w*\rho_\delta)(c_i*\rho_\delta)$ in (3.12) converge in $L^1(A)$ to $\partial\beta/\partial y_i(w)c_i$. Therefore, given an open set $A'\subset A$ and $\varphi\in C^\infty_c(A')$ with $|\varphi|\le 1$, by applying (3.7) with $Q=\operatorname{supp}\varphi$ we obtain

$$\begin{aligned} |\langle\sigma,\varphi\rangle| &= \lim_{\delta\downarrow 0}\langle\frac{\partial\beta(w*\rho_\delta)}{\partial t} + b\cdot\nabla\beta(w*\rho_\delta) - \sum_{i=1}^{d}\frac{\partial\beta}{\partial y_i}(w*\rho_\delta)c_i*\rho_\delta\mathscr{L}^{d+1},\varphi\rangle \\ &\le L\sup_{B_L}|\nabla\beta|\limsup_{\delta\downarrow 0}\int_Q |r_\delta|\,dt\,dx \le L\sup_{B_L}|\nabla\beta|I(\rho)|D^sb|(A'). \end{aligned}$$

Since φ is arbitrary we obtain that $|\sigma|(A')\le L\sup_{B_L}|\nabla\beta|I(\rho)|D^sb|(A')$ for any open set $A'\subset A$, and therefore $|\sigma|\le L\sup_{B_L}|\nabla\beta|I(\rho)|D^sb|$ as measures in A. This shows that σ is singular with respect to $\mathscr{L}^{d+1}$.
STEP 3. In this step we show that the measure σ in Step 2 vanishes on A. To this aim, we fix again an open set $A'\subset A$ and $\varphi\in C^\infty_c(A')$ with $|\varphi|\le 1$, and apply this time (3.6) instead of (3.7), with $Q=\operatorname{supp}\varphi$, to obtain by the same argument of Step 2

$$|\langle\sigma,\varphi\rangle| \le L\sup_{B_L}|\nabla\beta|\int_A \Lambda(M_t,\rho)\,d|D^sb| + L\sup_{B_L}|\nabla\beta|(d+I(\rho))|D^ab|(A).$$

As φ is arbitrary this leads to the inequality between measures in A

$$|\sigma| \le L\sup_{B_L}|\nabla\beta|\Lambda(M_t,\rho)\,d|D^sb| + L\sup_{B_L}|\nabla\beta|(d+I(\rho))|D^ab|,$$

and since we already know that $\sigma\perp\mathscr{L}^{d+1}$ and $|D^ab|\ll\mathscr{L}^{d+1}$ we infer

$$|\sigma|\le L'\Lambda(M_t,\rho)\,d|D^sb| \qquad \text{in } A, \text{ with } L' := L\sup_{B_L}|\nabla\beta|.$$

Notice that σ does not depend on the convolution kernel ρ, so that we can optimize this estimate as follows: let g be the Radon–Nikodym derivative of $|\sigma|$ with respect to $|D^sb|$ and notice that the inequality above gives $g\le L'\Lambda(M_t,\rho)$ $|D^sb|$-a.e. for any convolution kernel ρ with compact support. Now, let $D\subset C^\infty_c(B_1)$ be a contable dense set, with respect to the $W^{1,1}(B_1)$ norm, in the set

$$R := \left\{\rho\in W^{1,1}(B_1):\ \rho\ge 0,\ \int_{B_1}\rho = 1\right\}.$$

Since D is countable we have $g \leq L' \inf_{\rho \in D} \Lambda(M_t, \rho)$ $|D^s b|$-a.e. and since D is dense we obtain

$$g(t,x) \leq L' \inf_{\rho \in R} \Lambda(M_t(x), \rho) \qquad \text{for } |D^s b|\text{-a.e. } (t,x) \in A.$$

Taking into account the rank one structure of M_t given by (3.5), Lemma 3.3 gives that $g = 0$ $|D^s b|$-a.e. Therefore $|\sigma| = 0$. $\square$

If we suppress the time dependence then our renormalization result reads as follows:

Corollary 3.6. *Let $\Omega \subset \mathbb{R}^d$ be open and let $B \in BV_{\mathrm{loc}}(\Omega; \mathbb{R}^d)$, $w \in L^\infty_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$. Assume that*

$$B \cdot \nabla w_i = c_i \mathscr{L}^d \quad i = 1, \ldots, m \qquad \textit{and} \qquad D \cdot B \ll \mathscr{L}^d.$$

Then $B \cdot \nabla \beta(w) = \sum_{i=1}^{m} \frac{\partial \beta}{\partial y_i}(w) c_i \mathscr{L}^d$ for any C^1 function $\beta : \mathbb{R}^m \to \mathbb{R}$.

Remark 3.7 (How the use of the rank-one theorem could be avoided). Since Alberti's rank one theorem is a non-trivial measure theoretic result, one might ask whether its role in our proof is really essential or not (in this respect, notice that the vector field considered in [8] trivially satisfies the rank-one property). Using the fact that $\int_{\mathbb{R}^d} z_i \partial\rho/\partial z_j \, dz = -\delta_{ij}$ for any convolution kernel ρ we obtain

$$\Lambda(M, \rho) \geq \left| \int \langle Mz, \nabla \rho(z) \rangle \, dz \right| = |\mathrm{trace}\,(M)|. \tag{3.13}$$

Therefore a necessary condition for the infimum in (3.10) being 0 is that $\mathrm{trace}\,(M) = 0$. Alberti recently showed [5] that this condition is also sufficient, but the construction of the (asymptotically) optimal kernels ρ is much more delicate, compared to the simple one given in Lemma 3.3 for the case when M is a rank one matrix with 0 divergence. For this reason, although in the light of this result the only essential information for our construction is that $\mathrm{trace}\,(M_t(x)) = 0$ (coming from the fact that $D \cdot b_t \ll \mathscr{L}^d$), we kept our original proof of the renormalization property which uses the rank one theorem.

In the following remark we point out some extensions of our result.

Remark 3.8. (1) Singular divergence. The counterexamples in Section IV of [26] indicate the necessity of bounds on the divergence with respect to the spatial variables (in particular for the validity of comparison results, see the next section). In the BV case our proof fails if $D_x \cdot b(t, \cdot)$ is not absolutely continuous with respect to $\mathscr{L}^d$: indeed, in this case the vectors ξ_t, η_t in (3.5) are not orthogonal in sets with positive measure and the lower bound (3.13) shows that there is a smaller gain in choosing anisotropic smoothing kernels. However, our proof can be adapted to show that

$$\frac{\partial \beta(w)}{\partial t} + b \cdot \nabla \beta(w) = \sum_{i=1}^{m} c_i \frac{\partial \beta}{\partial y_i}(w) \mathscr{L}^{d+1} + \sigma_\beta \qquad \text{in } \mathcal{D}'((0,T) \times \Omega)$$

where σ_β satisfies

$$|\sigma_\beta|\llcorner A \le \|w\|_{L^\infty(A)} \sup_{|y|\le\|w\|_{L^\infty(A)}} |\nabla\beta(y)| \int_0^T |\langle\eta_t,\xi_t\rangle||D^s b_t|\,dt$$

for any open set $A \subset\subset (0,T)\times\Omega$, with $M_t = \eta_t\otimes\xi_t$ (given by Alberti's theorem). The representation of this "defect measure" σ_β is an open problem, related to the analysis done in the recent papers [7], [23], [24].
(2) Vector fields with singularities. As in [19], in Corollary 3.6 the $BV_{\rm loc}$ assumption on B can be made only in an open set $\Omega_0\subset\Omega$ with $\mathscr{H}^{d-1}(\Omega\setminus\Omega_0)=0$, provided $w\in L^\infty_{\rm loc}(\Omega)$ and $B\in L^\infty_{\rm loc}(\Omega;\mathbb{R}^d)$. Indeed, under these assumptions the singularity $\Omega\setminus\Omega_0$ is removable by Lemma 4.1 of [19].
(3) Measurable, BV and $W^{1,1}$ dependence. If $A\subset\mathbb{R}^m$ is an open set, the method we used to show Theorem 3.5 could perhaps be used for vector fields

$$B(t,x) = (a(t,x), b(t,x))\,,\quad a: A\times\Omega\to\mathbb{R}^m,\quad b: A\times\Omega\to\mathbb{R}^d$$

assuming the following regularity of a and b with respect to the "time" and "space" variables (t,x):

$$\begin{pmatrix} W^{1,1} & \text{measurable} \\ \text{measurable} & BV \end{pmatrix}$$

with suitable local uniform bounds as in (3.2) and with absolute continuity of the distributional divergence. But surely this extension requires additional technical complications (since mollifying only with respect to the x variable does not produce anymore $W^{1,1}$ functions in (t,x)) and it is beyond the main scopes of this paper.

4 Uniqueness and comparison results for the transport equation

In this section we apply the renormalization property to derive uniqueness and comparison results for solutions $w(t,x)$ of the transport equation. In this section we assume that $b(t,x):(0,T)\times\mathbb{R}^d\to\mathbb{R}^d$ satisfies the conditions (3.1) with $\Omega=\mathbb{R}^d$ and

$$\int_{(0,T)\times B_R} |b_t|\,dt\,dx + \int_I |Db_t|(B_R)\,dt < +\infty \qquad \text{whenever } R>0 \text{ and } \overline{I}\subset(0,T). \tag{4.1}$$

We assume also that for $\mathscr{L}^1$-a.e. $t\in(0,T)$ the distribution $D\cdot b_t$ is representable as $\operatorname{div} b_t\mathscr{L}^d$ with

$$\int_0^T \|\operatorname{div} b_t\|_{L^\infty(B_R)}\,dt < +\infty \qquad \text{whenever } R>0. \tag{4.2}$$

Given $w : (0,T) \times \mathbb{R}^d \to \mathbb{R}$ satisfying $w \in L^\infty((0,T) \times B_R)$ for any $R > 0$, for notational simplicity we set $w_t(x) = w(t,x)$. Choosing $c = w_t e_t$, with $e_t(x) = e(t,x)$, the PDE studied in the previous section becomes

$$\frac{\partial w_t}{\partial t} + b_t \cdot \nabla w_t = e_t w_t \qquad \text{in } (0,T) \times \mathbb{R}^d. \tag{4.3}$$

It is immediate to check that (4.3) is equivalent to

$$\frac{d}{dt}\int_{\mathbb{R}^d} w_t \varphi \, dx = \int_{\mathbb{R}^d} (e_t + \operatorname{div} b_t) w_t \varphi \, dx + \int_{\mathbb{R}^d} w_t b_t \cdot \nabla \varphi \, dx \qquad \text{in } \mathcal{D}'((0,T))$$

for any $\varphi \in C_c^\infty(\mathbb{R}^d)$. Therefore, if

$$\int_0^T \left(\|e_t + \operatorname{div} b_t\|_{L^1(B_R)} \right) \|w_t\|_{L^\infty(B_R)} \, dt < +\infty \qquad \text{for any } R > 0$$

all maps $t \mapsto \int w_t \varphi \, dx$ are $W^{1,1}$ in $(0,T)$. Possibly modifying w_t in a $\mathscr{L}^1$-neglibile set of times we obtain that $t \mapsto w_t$ is continuous with respect to the $w^* - L^\infty(B_R)$ topology for any $R > 0$. Moreover there exist unique maps w_0, $w_T \in L^\infty_{\mathrm{loc}}(\mathbb{R}^d)$ such that

$$\lim_{t \downarrow 0} \int_{\mathbb{R}^d} w_t \varphi \, dx = \int_{\mathbb{R}^d} w_0 \varphi \, dx, \qquad \lim_{t \uparrow T} \int_{\mathbb{R}^d} w_t \varphi \, dx = \int_{\mathbb{R}^d} w_T \varphi \, dx$$

for any $\varphi \in C_c^\infty(\mathbb{R}^d)$.

Finally, it will be useful to extend the PDE also for negative times. This can be achieved setting

$$b_t(x) = 0, \quad e_t(x) = 0, \quad w_t(x) = w_0(x) \qquad \text{for } t < 0, \tag{4.4}$$

so that still $D \cdot b_t \ll \mathscr{L}^{d+1}$ for $\mathscr{L}^1$-a.e. $t \in (-\infty, T)$ and

$$\frac{d}{dt}\int_{\mathbb{R}^d} w_t \varphi \, dx = \int_{\mathbb{R}^d} (e_t + \operatorname{div} b_t) w_t \varphi \, dx + \int_{\mathbb{R}^d} w_t b_t \cdot \nabla \varphi \, dx \qquad \text{in } \mathcal{D}'((-\infty,T))$$

for any $\varphi \in C_c^\infty(\mathbb{R}^d)$. Equivalently, the transport equation

$$\frac{\partial w_t}{\partial t} + b_t \cdot \nabla w_t = e_t w_t \qquad \text{in } (-\infty,T) \times \mathbb{R}^d \tag{4.5}$$

holds.

Theorem 4.1. *Let w_t^i, $i = 1, 2$, be solutions of the transport equation (4.3) in $(0,T) \times \mathbb{R}^d$ with $\int_0^T \|e_t\|_{L^\infty(B_R)} \, dt < +\infty$ for any $R > 0$. Assume that b satisfies (3.1) with $\Omega = \mathbb{R}^d$, (4.1), (4.2) and that there exist constants $C \geq 0$, $R > 0$ such that*

$$\|w_t^i\|_\infty \leq C, \quad w_t^i = 0 \ \mathscr{L}^d\text{-a.e. in } \mathbb{R}^d \setminus B_R \tag{4.6}$$

for any $t \in (0,T)$. Then $w_0^1 \leq w_0^2$ implies $w_t^1 \leq w_t^2$ for any $t \in [0,T]$.

Proof. We extend b_t, e_t, w^1_t, w^2_t for negative times as in (4.4) to obtain (4.5) with $w_t = w^i_t$.

Let $\beta(t) = \sqrt{1+(t^+)^2} - 1 \in C^1(\mathbb{R})$. By Theorem 3.5 with $w_t = w^1_t - w^2_t$ we obtain

$$\frac{\partial \beta(w_t)}{\partial t} + b_t \cdot \nabla \beta(w_t) = e_t \beta'(w_t) w_t \mathscr{L}^{d+1} \qquad \text{in } (-\infty, T) \times \mathbb{R}^d,$$

therefore, using the inequalities $0 \le s\beta'(s) \le 2\beta(s)$ and (4.6), we get

$$\begin{aligned} \frac{d}{dt}\int_{\mathbb{R}^d} \beta(w_t)\,dx &= \int_{\mathbb{R}^d} e_t \beta'(w_t) w_t\,dx + \int_{\mathbb{R}^d} \operatorname{div} b_t \beta(w_t)\,dx \\ &\le \left(2\|e_t\|_{L^\infty(B_R)} + \|\operatorname{div} b_t\|_{L^\infty(B_R)}\right) \int_{\mathbb{R}^d} \beta(w_t)\,dx \end{aligned}$$

in the distribution sense in $(-\infty, T)$. Since $\beta(w_t) = 0$ for $t < 0$, from Gronwall Lemma we obtain that $\beta(w_t) = 0$ for any $t \in (-\infty, T]$. □

Remark 4.2. In the case when $e_t = -\operatorname{div} b_t$, studied in the next sections, we may use the identity

$$\frac{d}{dt}\int_{\mathbb{R}^d} \beta(w_t)\,dx = \int_{\mathbb{R}^d} \operatorname{div} b_t(\beta(w_t) - \beta'(w_t) w_t)\,dx$$

and the fact that $-\beta(t) \le \beta(t) - t\beta'(t) \le 0$ (due to the convexity of the function β used in the proof of Theorem 4.1) to replace the bounds on e_t and $\operatorname{div} b_t$ by

$$\int_0^T \|[\operatorname{div} b_t]^-\|_{L^\infty(B_R)}\,dt < +\infty \qquad \forall R > 0. \tag{4.7}$$

5 Probability measures in $C([0,T];\mathbb{R}^m)$

Let $A \subset \mathbb{R}^d$ be a $\mathscr{L}^d$-measurable set and a complete and separable metric space X. We recall that a measure-valued map $\eta_x : A \to \mathscr{M}_1(X)$ is said to be $\mathscr{L}^d$-measurable if $x \mapsto \eta_x(C)$ is $\mathscr{L}^d$-measurable for any Borel set $C \subset X$. If X is also locally compact, viewing $\mathscr{M}_1(X)$ as the unit ball of $(C_c(X))^*$ endowed with the weak* topology, it is not hard to see that measurability in this sense is equivalent to measurability in the conventional sense, where the σ-algebra in the codomain is the Borel σ-algebra.

The following theorem provides a simple necessary and sufficient condition ensuring that a family of probability measures γ_x in $C([0,T];\mathbb{R}^m)$ is made by Dirac masses. Notice that is easy to check that $\gamma \in \mathscr{P}(C([0,T];\mathbb{R}^m))$ is a Dirac mass if and only if η_t is a Dirac mass for any $t \in [0,T]$, where η_t is the image of γ under the map $\psi \mapsto \psi(t)$. Since the maps ψ are continuous, it suffices actually to check this condition only for a dense set of times.

Theorem 5.1. *Let $A \subset \mathbb{R}^d$ be a $\mathscr{L}^d$-measurable set, let $\gamma_x : A \to \mathscr{P}(C([0,T];\mathbb{R}^m))$ be a $\mathscr{L}^d$-measurable family and for any $t \in [0,T]$ define η_{tx} as the push forward of γ_x under the map $\psi \mapsto \psi(t)$. Assume that*

$$\bigcup_{x \in A} \operatorname{supp} \eta_{tx} \subset\subset \mathbb{R}^m \tag{5.1}$$

and that

$$\gamma_x\left(\{\psi : \ \psi(t) \in K\}\right) \gamma_x\left(\{\psi : \ \psi(t) \in K'\}\right) = 0 \ \textit{for}\ \mathscr{L}^d\textit{-a.e.}\ x \in A \tag{5.2}$$

whenever K, $K' \in \mathcal{B}(\mathbb{R}^m)$ are disjoint, for any $t \in [0,T]$.
Then γ_x is a Dirac mass for $\mathscr{L}^d$-a.e. $x \in A$.

Proof. Using the elementary criterion recalled before the statement of the theorem, it suffices to check that for any $t \in [0,T]$ we have that η_{tx} is a Dirac mass for $\mathscr{L}^d$-a.e. $x \in A$, where η_{tx} is the image of the measure γ_x under the map $\psi \mapsto \psi(t)$.

We can assume with no loss of generality that the supports of all measures η_x are contained in the closed unit cube Q and we consider the canonical decomposition of Q in closed 2^{mh} dyadic cubes Q_i^h with side length 2^{-h}.

Let us split Q in two parts, say Q' and Q''. By applying (5.2) we know that there are disjoint Borel sets A' and A'' whose union covers $\mathscr{L}^d$-almost all of A such that $\psi(t) \in Q'$ for γ_x-a.e. ψ for any $x \in A'$ and $\psi(t) \in Q''$ for γ_x-a.e. ψ for any $x \in A''$. Fix now h and repeat this argument finitely many times to obtain pairwise disjoint Borel sets A_i^h whose union covers $\mathscr{L}^d$-almost all of A and such that $\psi(t) \in Q_i^h$ for γ_x-a.e. ψ for any $x \in A_i^h$.

Fix i and a point of density 1 for A_i^h that is also a Lebesgue point for $y \mapsto \eta_{ty}(\mathbb{R}^m \setminus Q_i^h)$. Since this map vanishes on A_i^h we obtain that $\eta_{tx}(\mathbb{R}^m \setminus Q_i^h) = 0$, i.e., $\operatorname{supp} \eta_{tx} \subset Q_i^h$. This proves that for any h the diameter of the support of η_x is less than $\sqrt{m}2^{-h}$ for $\mathscr{L}^d$-a.e. $x \in A$. Letting $h \to \infty$ the proof is achieved. □

We fix now a bounded vector field $b : (0,T) \times \mathbb{R}^d \to \mathbb{R}^d$ and define the space $\mathscr{S}(\mathbb{R}^d)$ of continuous $\mathbb{R}^d$-valued maps in $[0,T]$ and the subspace $\mathscr{S}_x^b(\mathbb{R}^d)$ made up by solutions of the ODE $\dot\psi = b(t,\psi)$ starting at x.

Definition 5.2 (Spaces of continuous maps). We set $\mathscr{S}(\mathbb{R}^d) = C([0,T];\mathbb{R}^d)$, endowed with the sup norm. Given $x \in \mathbb{R}^d$ we denote by $\mathscr{S}_x^b(\mathbb{R}^d)$ the subspace

$$\mathscr{S}_x^b(\mathbb{R}^d) := \left\{ \psi \in \mathscr{S}(\mathbb{R}^d) : \ \psi(t) = x + \int_0^t b(\tau,\psi(\tau))\,d\tau \quad \forall t \in [0,T] \right\}.$$

Notice that at this level of generality the definition of $\mathscr{S}_x^b(\mathbb{R}^d)$ is sensitive to modifications of b even in very small sets. Therefore b is a given function and should be not be thought as an equivalence class (but see the comments after Definition 6.1).

Next, we show how a family γ_x of probability measures in $\mathscr{S}(\mathbb{R}^d)$ concentrated on $\mathscr{S}^b_x(\mathbb{R}^d)$ induces a solution of the transport equation. Let $A \subset \mathbb{R}^d$ be a $\mathscr{L}^d$-measurable set and let $\gamma_x : A \to \mathscr{M}_1(\mathscr{S}(\mathbb{R}^d))$ be a $\mathscr{L}^d$-measurable map. The measures

$$\mu_t^{\gamma_x}(\varphi) := \int_A \int_{\mathscr{S}(\mathbb{R}^d)} \varphi(\psi(t))\, d\gamma_x(\psi)\, dx \qquad \varphi \in C_c(\mathbb{R}^d), \quad t \in [0,T] \tag{5.3}$$

are well defined in $\mathbb{R}^d$ provided $\int_A \gamma_x(\mathscr{S}(\mathbb{R}^d))\, dx < +\infty$.

Theorem 5.3. *Assume that b is locally bounded in $(0,T)\times\mathbb{R}^d$ and $\gamma_x\left(\mathscr{S}(\mathbb{R}^d) \setminus \mathscr{S}^b_x(\mathbb{R}^d)\right) = 0$ for $\mathscr{L}^d$-a.e. $x \in A$. Then the measures $\mu_t = \mu_t^{\gamma_x}$ defined above solve the transport equation*

$$\frac{\partial \mu_t}{\partial t} + D_x \cdot (b\mu_t) = 0 \qquad \textit{in } \mathcal{D}'((0,T)\times\mathbb{R}^d)$$

and $\mu_0 = \gamma_x(\mathscr{S}(\mathbb{R}^d))\mathscr{L}^d \llcorner A$. In particular, if $D \cdot b_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$ and $\mu_t = w_t \mathscr{L}^d$ with $w_t(x) \in L^\infty_{\rm loc}((0,T)\times\mathbb{R}^d)$ we have

$$\frac{\partial w_t}{\partial t} + b_t \cdot \nabla w_t = -w_t \mathrm{div}\, b_t \qquad \textit{in } \mathcal{D}'((0,T)\times\mathbb{R}^d). \tag{5.4}$$

Proof. Let λ be the finite measure in $A \times \mathscr{S}(\mathbb{R}^d)$ whose disintegration with respect to the x variable is γ_x, i.e.

$$\int_{A\times\mathscr{S}(\mathbb{R}^d)} g(x,\psi)\, d\lambda(x,\psi) = \int_A \int_{\mathscr{S}(\mathbb{R}^d)} g(x,\psi)\, d\gamma_x(\psi)\, dx$$

for any bounded Borel function g. Then, we know that for λ-a.e. (x,ψ) the identity $\dot\psi(t) = b(\psi(t))$ is true for $\mathscr{L}^1$-a.e. $t \in (0,T)$. By Fubini's theorem we have

$$\dot\psi(t) = b(t,\psi(t)) \qquad \text{for } \lambda\text{-a.e. } (x,\psi) \in A \times \mathscr{S}(\mathbb{R}^d)$$

for $\mathscr{L}^1$-a.e. $t \in (0,T)$. Therefore, for any $\varphi \in C^\infty_c(\mathbb{R}^d)$ we have

$$\begin{aligned} \frac{d}{dt}\mu_t(\varphi) &= \frac{d}{dt}\int_A \int_{\mathscr{S}(\mathbb{R}^d)} \varphi(\psi(t))\, d\lambda(x,\psi) = \int_{\mathbb{R}^d}\int_{\mathscr{S}(\mathbb{R}^d)} \nabla\varphi(\psi(t)) \cdot b_t(\psi(t))\, d\lambda(x,\psi) \\ &= \int_{\mathbb{R}^d} b_t \cdot \nabla\varphi\, d\mu_t \qquad \text{in } \mathcal{D}'((0,T)) \end{aligned}$$

and the transport equation holds. □

The following Theorem is, together with Theorem 3.5, one of the main results of this paper. It provides a sufficient condition ensuring that the measures γ_x concentrated in $\mathscr{S}^b_x(\mathbb{R}^d) \subset \mathscr{S}(\mathbb{R}^d)$ are actually Dirac masses. In the proof we use the fact that, by a localization argument in the ψ variable, the comparison principle of Theorem 4.1 shows that the property of being orthogonal for measures (or, rather, functions) is preserved under the flow (4.3).

Theorem 5.4. *Let $A \subset \mathbb{R}^d$ be a $\mathscr{L}^d$-measurable set, let $\gamma_x : A \to \mathscr{P}(\mathscr{S}(\mathbb{R}^d))$ be a $\mathscr{L}^d$-measurable family of probability measures satisfying:*

(i) $\gamma_x \left(\mathscr{S}(\mathbb{R}^d) \setminus \mathscr{S}^b_x(\mathbb{R}^d)\right) = 0$ *for* $\mathscr{L}^d$*-a.e.* $x \in A$.

(ii) The measures $\mu_t^{\gamma_x}$ *in (5.3) are absolutely continuous with respect to* $\mathscr{L}^d$ *and their densities* $w_t(y)$ *are bounded by a constant independent of* $y \in \mathbb{R}^d$ *and* $t \in (0,T)$.

(iii) The vector field b *satisfies (3.1) with* $\Omega = \mathbb{R}^d$*, (4.1) and (4.7).*

Then γ_x *is a Dirac mass for* $\mathscr{L}^d$*-a.e.* $x \in A$.

Proof. We fix $t \in [0,T]$ and denote by η_{tx} be the push forward of γ_x under the map $\psi \mapsto \psi(t)$. We will check that the assumption (5.2) in Theorem 5.1 holds. If not, we can find $t \in (0,T]$ and disjoint Borel sets K, $K' \subset \mathbb{R}^d$ and a Borel set C with $\mathscr{L}^d(C) > 0$ such that

$$\gamma_x\left(\{\psi :\ \psi(t) \in K\}\right) \gamma_x\left(\{\psi :\ \psi(t) \in K'\}\right) > 0 \qquad \forall x \in C.$$

Possibly passing to a smaller set still with positive Lebesgue measure we can assume that

$$0 < \gamma_x(\{\psi(t) \in K\}) \leq M\gamma_x(\{\psi(t) \in K'\}).$$

for some constant M. We define

$$\gamma_x^1 := \chi_C(x)\gamma_x \llcorner \{\psi :\ \psi(t) \in K\}, \qquad \gamma_x^2 := M\chi_C(x)\gamma_x \llcorner \{\psi :\ \psi(t) \in K'\}$$

and denote by μ_t^i the measures associated to γ_x^i as in (5.3). Then, by Theorem 5.3 the measures μ_t^i solve the transport equation (5.4) and

$$\mu_0^1 = \eta_{tx}(K)\mathscr{L}^d \llcorner C, \qquad \mu_0^2 = M\eta_{tx}(K')\mathscr{L}^d \llcorner C,$$

so that $\mu_0^1 \leq \mu_0^2$. On the other hand

$$\mu_t^1 = \int_C \eta_{tx} \llcorner K \, dx \perp M \int_C \eta_{tx} \llcorner K' \, dx = \mu_t^2.$$

Notice also that the supports μ_t^i are equi-bounded and assumption (ii) ensures that the densities of μ_t^i are uniformly bounded in space and time. This contradicts Theorem 4.1 and Remark 4.2. □

We have seen that L^∞ estimates on $[\operatorname{div} b_t]^-$ imply that the generalized solutions of the transport equation in (5.3) are actually classical. We show now that L^∞ estimates on $[\operatorname{div} b_t]^+$ imply that the flow $x \mapsto \gamma_x$ is injective in a weak sense.

In the proof of this fact we will need the following decomposition result, loosely related to the existence of conditional probability measures in Probability. The statement involves a $\mathscr{L}^d$-measurable map $x \mapsto K(x)$ with values in the compact subsets of $\mathbb{R}^m$; we denote by $\mathscr{K}(\mathbb{R}^m)$ the locally compact metric space of these compact subsets, endowed with the Hausdorff metric:

$$d_{\mathscr{H}}(K,L) := \inf\left\{r>0: \ K \subset L_r \quad \text{and} \quad L \subset K_r\right\}.$$

Here, as usual, K_r and L_r denote the open r-neighbourhoods of K and L.

Theorem 5.5. *Let $A \subset \mathbb{R}^d$ be a bounded $\mathscr{L}^d$-measurable set, let $\eta_x : A \to \mathscr{M}_1(\mathbb{R}^m)$ be a $\mathscr{L}^d$-measurable family satisfying*

$$\int_C \eta_x\,dx \perp \int_{C'} \eta_x\,dx \quad \text{whenever } C,\, C' \in \mathcal{B}(A) \text{ and } C \cap C' = \emptyset.$$

Then, for any $\delta > 0$ there exist a Borel set $C \subset A$ and, for $x \in C$, a $\mathscr{L}^d$-measurable family of pairwise disjoint compact sets $K(x)$ such that $\mathscr{L}^d(A \setminus C) < \delta$ and

$$\eta_x\left(\mathbb{R}^m \setminus K(x)\right) < \delta \qquad \forall x \in C.$$

Proof. Possibly embedding A in a larger cube Q we can assume that $A = Q$ is a cube. We first recall that in this context Lebesgue differentiation theorem reads as follows: for $\mathscr{L}^d$-a.e. $x \in Q$ we have

$$\lim_{i\to\infty} \frac{1}{\mathscr{L}^d(Q_i)} \int_{Q_i} \eta_y\,dy \quad \text{weakly converge to } \eta_x \tag{5.5}$$

for any family of cubes $Q_i \ni x$ with $\operatorname{diam}(Q_i) \to 0$. This can be easily deduced by applying the standard Lebesgue differentiation theorem to the functions $x \mapsto \int \varphi\,d\eta_x$, with $\varphi \in C_c(\mathbb{R}^m)$ varying in a countable dense set.

Without loss of generality we can assume that $Q = Q_1$ is the unit cube centered at the origin. We consider the canonical dyadic decomposition of Q in closed cubes Q_i^h, $h \in \mathbb{N}$ and $1 \le i \le 2^{hd}$ and the corresponding family of measures $\eta_i^h = \int_{Q_i^h} \eta_x\,dx$. We fix a sequence $(\delta_h) \subset (0,+\infty)$ with $\sum_1^\infty \delta_h < \delta$ and, by induction on h, we define a family of compact sets $K_i^h \subset \mathbb{R}^m$, $1 \le i \le 2^{hd}$, and countable families $\mathcal{F}_h$ of dyadic cubes of Q with side length at most 2^{-h} with the following properties:
(a) If $Q_i^h \subset Q_j^k$ then $K_i^h \subset K_j^k$;
(b) for any $h \ge 1$ the sets K_i^h, $1 \le i \le 2^{hd}$, are pairwise disjoint;
(c) the Lebesgue measure of the union of the cubes in $\mathcal{F}_h$ is less than δ_h;
(d) for $h > k \ge -1$ we have $\eta_j^h\left(K_l^{k-1} \setminus K_i^k\right) \le \delta_{k+1} 2^{-hd}$, whenever $Q_j^h \subset Q_i^k \subset Q_l^{k-1}$ and $Q_j^h \notin \mathcal{F}_{k+1}$ (with the convention $K_l^{-1} = \mathbb{R}^m$ and $Q_l^{-1} = \mathbb{R}^d$).

To begin with, we set $\mathcal{F}_0 = \emptyset$ and we choose K_1^0 such that $\eta_1^0(\mathbb{R}^m \setminus K_1^0) \le \delta_1^2$. Then, for $h \ge 1$ we say that a cube Q_i^h belongs to $\mathcal{F}_1$ if $\eta_i^h(\mathbb{R}^m \setminus K_1^0) \ge \delta_1 2^{-hd}$. If we choose a disjoint subfamily $Q_{i(l)}^{h(l)}$ of $\mathcal{F}_1$ whose union is the union of all cubes in $\mathcal{F}_1$, we have

$$\mathscr{L}^d\left(\bigcup_{Q_i^h\in\mathcal{F}_1} Q_i^h\right) \le \sum_l \mathscr{L}^d(Q_{i(l)}^{h(l)}) \le \frac{\eta_1^0(\mathbb{R}^m\setminus K_1^0)}{\delta_1} \le \delta_1.$$

Next we consider all cubes Q_j^1 and we repeat this construction inside all of them, finding pairwise disjoint compact sets $K_j^1 \subset K_1^0$ such that $\sum_j \eta_j^1(K_1^0\setminus K_j^1) \le \delta_2^2$ (this is possible because the measures η_j^1 are mutually orthogonal) and defining $\mathcal{F}_2$ as the collection of all cubes $Q_i^h \subset Q_j^1$ such that $h \ge 2$ and $\eta_i^h(K_1^0\setminus K_j^1) \ge \delta_2 2^{-hd}$. Again, if we choose a disjoint subfamily $Q_{i(l)}^{h(l)}$ of $\mathcal{F}_2$ whose union is the union of all cubes in $\mathcal{F}_2$, we have

$$\mathscr{L}^d\left(\bigcup_{Q_i^h\in\mathcal{F}_2} Q_i^h\right) \le \sum_l \mathscr{L}^d(Q_{i(l)}^{h(l)}) \le \sum_{j=1}^{2^d} \frac{\eta_j^1(K_1^0\setminus K_j^1)}{\delta_2} \le \delta_2.$$

It is easy to check that iterating this procedure all conditions (a)-(d) are satisfied.

We define C as the complement of the union of all cubes in the families $\mathcal{F}_h$. By condition (c) and our choice of δ_h we have $\mathscr{L}^d(C) < \delta$. We fix $x \in C$ where (5.5) holds and the coordinates x_i of x are not dyadic numbers (these condition are satisfied for $\mathscr{L}^d$-a.e. $x \in C$). Then x belongs to the intersection of a unique family of nested cubes $Q_{i(h)}^h$ with $Q_{i(h)}^h \notin \mathcal{F}_k$ for any $k < h$. Using (d) we obtain

$$\eta_{i(h)}^h(\mathbb{R}^m\setminus K_{i(h_0)}^{h_0}) \le \sum_{l=-1}^{h_0-1} \eta_{i(h)}^h(K_{i(l)}^l\setminus K_{i(l+1)}^{l+1}) \le \sum_{l=1}^{\infty} \delta_l 2^{-hd}$$

for any $h > h_0 \ge 0$. Defining $K(x)$ as $\cap_h K_{i(h)}^h$ and taking into account that, by condition (a), $h \mapsto K_{i(h)}^h$ is decreasing, from the weak convergence of $2^{hd}\eta_{i(h)}^h$ to η_x we get

$$\eta_x(\mathbb{R}^m\setminus K_{i(h_0)}^{h_0}) \le \sum_{l=1}^{\infty} \delta_l < \delta.$$

Letting $h_0 \to \infty$ we obtain $\eta_x(\mathbb{R}^m\setminus K(x)) < \delta$ and by condition (b) the sets $K(x)$ are pairwise disjoint. Moreover $K(x)$ is the limit of piecewise constant maps, and therefore it is $\mathscr{L}^d$-measurable. $\square$

Remark 5.6. By a standard exhaustion argument we can obtain a $\mathscr{L}^d$-negligible set $C \subset \mathbb{R}^d$ and, for $x \in \mathbb{R}^d\setminus C$, pairwise disjoint Borel sets $A(x)$ such that $\eta_x(A(x)) = 1$. Given C_i, K_i for $\delta = 2^{-i}$, we simply define $C = \limsup_i C_i$ and $A(x) = \liminf_i K_i(x)$ for any $x \in \mathbb{R}^d\setminus C$.

Theorem 5.7 (Weak injectivity of the flow). *Let $A \subset \mathbb{R}^d$ be a $\mathscr{L}^d$-measurable set, let $\gamma_x : A \to \mathscr{P}(\mathscr{S}(\mathbb{R}^d))$ be a $\mathscr{L}^d$-measurable family of probability measures satisfying:*

(i) $\gamma_x \left(\mathscr{S}(\mathbb{R}^d) \setminus \mathscr{S}^b_x(\mathbb{R}^d) \right) = 0$ *for $\mathscr{L}^d$-a.e. $x \in A$.*

(ii) The measures $\mu_t^{\gamma_x}$ in (5.3) are absolutely continuous with respect to $\mathscr{L}^d$ and their densities $w_t(y)$ are bounded by a constant independent of $y \in \mathbb{R}^d$ and $t \in (0,T)$.

(iii) The vector field b satisfies (3.1) with $\Omega = \mathbb{R}^d$, (4.1) and

$$\int_0^T \|[\mathrm{div}\, b_t]^+\|_{L^\infty(B_R)}\, dt < +\infty \qquad \forall R > 0. \tag{5.6}$$

Then, for any $t \in [0,T]$ we can find a $\mathscr{L}^d$-negligible set $N \subset A$ and, for $x \in A \setminus N$, pairwise disjoint σ-compact sets $A(x)$ such that $\eta_{tx}(A(x)) = 1$ for any $x \in A \setminus N$.

Proof. We check that the orthogonality condition in Theorem 5.5 is satisfied. Let C_1, $C_2 \subset A$ be disjoint Borel sets with $C_1 \cap C_2 = \emptyset$. If $\int_{C_1} \eta_{tx}\, dx$ is not orthogonal to $\int_{C_2} \eta_{tx}\, dx$ we can find a constant M and a Borel set $E \subset \mathbb{R}^d$ such that

$$\int_{C_1} \eta_{tx} \llcorner E\, dx \leq M \int_{C_2} \eta_{tx} \llcorner E\, dx$$

and $\int_{C_i} \eta_{tx}(E)\, dx > 0$. We define

$$\gamma^1_x := \chi_{C_1}(x) \gamma_x \llcorner \{\psi : \ \psi(t) \in E\}, \qquad \gamma^2_x := M \chi_{C_2}(x) \gamma_x \llcorner \{\psi : \ \psi(t) \in E\}$$

and denote by μ^i_t the measures associated to γ^i_x as in (5.3). Then, by Theorem 5.3 the measures μ^i_t solve the continuity equation (5.4), $\mu^1_0 = \eta_{tx}(E) \mathscr{L}^d \llcorner C_1$ and $\mu^2_0 = M \eta_{tx}(E) \mathscr{L}^d \llcorner C_2$ are orthogonal and

$$\mu^1_t = \int_{C_1} \eta_{tx} \llcorner E\, dx \leq M \int_{C_2} \eta_{tx} \llcorner E\, dx = \mu^2_t. \tag{5.7}$$

Notice also that the supports μ^i_t are equi-bounded (since b is bounded) and assumption (ii) ensures that the densities of μ^i_t are uniformly bounded in space and time. Then, the inequality (5.7) contradicts Theorem 4.1 and Remark 4.2, by reversing the time variable. By Remark 5.6 the conclusion follows. □

6 Lagrangian flows

In this section we define a suitable concept of (regular) Lagrangian flow for bounded and vector fields $b(t,x)$ having locally bounded variation with respect to x, showing natural existence, uniqueness, stability and injectivity results. The main tools in our analysis are Theorem 5.4 and the theory of Young measures.

Definition 6.1 (Lagrangian flows). Let $A \subset \mathbb{R}^d$ be a $\mathscr{L}^d$-measurable set, let $\psi : B \to \mathscr{S}(\mathbb{R}^d)$ be a $\mathscr{L}^d$-measurable map. We say that ψ is a *Lagrangian flow* starting from A, relative to b, if $\psi(x) \in \mathscr{S}^b_x(\mathbb{R}^d)$ for $\mathscr{L}^d$-a.e. $x \in A$.
We say that a Lagrangian flow ψ is *regular* if there exists an increasing union of $\mathscr{L}^d$-measurable sets A_h whose union is A and constants C_h such that

$$\left|\int_{A_h} \varphi\left(\psi(x)(t)\right)\, dx\right| \leq C_h \int_{\mathbb{R}^d} |\varphi(y)|\, dy \qquad \forall \varphi \in C_c(\mathbb{R}^d),\ t \in [0,T]. \tag{6.1}$$

Finally, we say that a regular Lagrangian flow ψ is (essentially) *injective* if

$$\int_{C_1} \delta_{\psi(x)(t)}\, dx \perp \int_{C_2} \delta_{\psi(x)(t)}\, dx \qquad \forall t \in [0,T]$$

for any pair of disjoint $\mathscr{L}^d$-measurable sets C_1, $C_2 \subset A$.

Clearly $\mathscr{S}^b_x(\mathbb{R}^d)$ can change dramatically if b is replaced by another field $\hat{b}$ in the same equivalence class (i.e. equal to b $\mathscr{L}^{d+1}$-a.e. in $(0,T) \times \mathbb{R}^d$). But the concept of regular Lagrangian flow is invariant under this transformation: indeed, if $N \subset (0,T) \times \mathbb{R}^d$ is any $\mathscr{L}^{d+1}$-negligible set Borel set, setting $N_t = \{x :\ (t,x) \in N\}$ from (6.1) we get

$$\int_A \chi_{N_t} \circ \psi(x)(t)\, dx = 0 \qquad \text{for any } t \text{ such that } \mathscr{L}^d(N_t) = 0,$$

so that, by integration with respect to t, Fubini's theorem gives

$$\int_0^T \chi_{N_t} \circ \psi(x)(t)\, dt = 0 \qquad \text{for } \mathscr{L}^d\text{-a.e. } x \in A. \tag{6.2}$$

Choosing $N = \{b \neq \hat{b}\}$ we see that the $\psi(x) \in \mathscr{S}^b_x(\mathbb{R}^d)$ for $\mathscr{L}^d$-a.e. $x \in A$ if and only if $\psi(x) \in \mathscr{S}^{\hat{b}}_x(\mathbb{R}^d)$ for $\mathscr{L}^d$-a.e. $x \in A$, i.e., regularity is invariant.

Theorem 6.2 (Existence). *Assume that*

(i) $b \in L^\infty((0,T) \times \mathbb{R}^d; \mathbb{R}^d)$;

(ii) $b_t \in BV_{\mathrm{loc}}(\mathbb{R}^d)$ *for* $\mathscr{L}^1$*-a.e.* $t \in (0,T)$, $|Db_t|(B_R) \in L^1_{\mathrm{loc}}(0,T)$ *for any* $R > 0$ *and*

$$\int_0^T \|[\operatorname{div} b_t]^-\|_{L^\infty(B_R)}\, dt < +\infty \qquad \forall R > 0.$$

Then, for any $\mathscr{L}^d$*-measurable set* $A \subset \mathbb{R}^d$ *there exists a regular Lagrangian flow starting from* A*. If* b *satisfies also (5.6) then the flow is injective.*

The proof of Theorem 6.2 will be obtained as a byproduct of the proof of the stability theorem Theorem 6.6, writing A as the union of a sequence of bounded subsets A_h, approximating b by convolution and using the following well known remark.

Remark 6.3 (Regularity of Lipschitz flows). Let $b(t,x)$ be bounded, $\mathscr{L}^d$-measurable with respect to t and uniformly (as t varies in $(0,T)$) Lipschitz continuous on B_R with respect to x for any $R > 0$. Then b induces a regular flow: indeed, the solutions of the ODE are unique and the Jacobian determinant $J(t,x)$ of the map $x \mapsto \psi(x)(t)$ satisfies (see for instance [26])

$$\frac{\partial J}{\partial t}(t,x) = \operatorname{div} b(t,\psi(x)(t))J(t,x) \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T),$$

for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$. Since $J(0,x) = 1$, by integration in time we obtain

$$e^{-C_R(t)} \le J(t,x) \le e^{D_R(t)} \quad \forall t \in [0,T], \qquad \text{for } \mathscr{L}^d\text{-a.e. } x \in B_R$$

with

$$C_R(t) := \int_0^t \|[\operatorname{div} b_t]^-\|_{L^\infty(B_{M_R})}\, dt, \qquad D_R(t) := \int_0^t \|[\operatorname{div} b_t]^+\|_{L^\infty(B_{M_R})}\, dt$$

and $M_R = R + T\|b\|_\infty$. In particular the area formula gives

$$\left|\int_{B_R} \varphi(\psi(x)(t))\, dx\right| \le e^{C_R(t)} \int_{B_R} |\varphi|(\psi(x)(t))J(t,x)\, dx \le e^{C_R(t)} \int_{B_{M_R}} |\varphi(y)|\, dy$$

for any $\varphi \in C^\infty_c(\mathbb{R}^d)$, so that (6.1) holds. Notice that this upper bound requires only an upper bound on the negative part of $\operatorname{div} b_t$.

Theorem 6.4 (Uniqueness). *Assume that (i) and (ii) in Theorem 6.2 hold. Then, if* ψ_i *are regular Lagrangian flows starting from* $\mathscr{L}^d$*-measurable set* $A_i \subset \mathbb{R}^d$*,* $i = 1, 2$*, we have*

$$\psi_1(x) = \psi_2(x) \quad \textit{for } \mathscr{L}^d\textit{-a.e. } x \in A_1 \cap A_2.$$

Proof. Using the validity of (6.1) for ψ_i we can assume with no loss of generality that A_i are bounded and that the measures

$$\mu_t^i(\varphi) := \int_{A_i} \varphi\left(\psi(x)(t)\right)\, dx, \quad \varphi \in C_c(\mathbb{R}^d), \qquad t \in [0,T]$$

are representable by $w_t^i \mathscr{L}^d$, with $\|w_t^i\|_\infty \le C$ independent of t and i. Assume by contradiction that

$$A := \{x \in A_1 \cap A_2 :\ \psi_1(x) \neq \psi_2(x)\}$$

has strictly positive Lebesgue measure. We define

$$\gamma_x := \frac{1}{2}\left(\delta_{\psi_1(x)} + \delta_{\psi_2(x)}\right).$$

Then γ_x is not a Dirac mass for any $x \in A$ and it is concentrated on $\mathscr{S}_x^b(\mathbb{R}^d)$ for $\mathscr{L}^d$-a.e. $x \in A$. By our initial assumption also the condition (ii) in Theorem 5.4 is fulfilled. This gives a contradiction with the conclusion of that theorem. $\square$

Finally, we show stability of regular Lagrangian flows with respect to approximation by locally Lipschitz and uniformly bounded vector fields b_h, assuming the uniform bound (6.6) on the flows. By Remark 6.3 this bound is fulfilled if

$$\sup_h \int_0^T \|[\mathrm{div}\,(b_h)_t]^-\|_{L^\infty(B_R)}\, dt < +\infty \qquad \forall R > 0.$$

We recall the following basic result on the theory of Young measures.

Theorem 6.5 (Fundamental theorem on Young measures). *Let K be a compact metric space, let $A \subset \mathbb{R}^d$ be a bounded $\mathscr{L}^d$-measurable set and let η_{hx} be a sequence of measure-valued maps from A to $\mathscr{P}(K)$. Then there exist a measure-valued map $\eta_x : A \to \mathscr{P}(K)$ and a subsequence $h(k)$ such that*

$$\lim_{k\to\infty} \int_A \int_K \phi(x,u)\, d\eta_{h(k)x}(u)\, dx = \int_A \int_K \phi(x,u)\, d\eta_x(u)\, dx$$

for any bounded function $\phi(x,u) : A \times K \to R$ continuous with respect to u and $\mathscr{L}^d$-measurable with respect to x.

Proof. See for instance [34] or §2 of [4]; in functional analytic terms it can be derived from the duality $(L^1(A; C(K))^* = L^\infty\left(A; (C(K))^*\right)$. $\square$

Theorem 6.6 (Stability). *Assume that (i) and (ii) in Theorem 6.2 hold. Let $b_h(t,x)$: $(0,T) \times \mathbb{R}^d \to \mathbb{R}^d$ be satisfying*

$$b_h(\cdot, x) \text{ is } \mathscr{L}^1\text{-measurable in } (0,T) \text{ for any } x \in \mathbb{R}^d, \tag{6.3}$$

$$\sup_h \|b_h\|_\infty \le C \quad \textit{and} \quad b_h \to b \ \textit{in} \ L^1_{\rm loc}((0,T)\times\mathbb{R}^d;\mathbb{R}^d), \tag{6.4}$$

$$\nabla b_h(t,\cdot) \in L^\infty((0,T)\times B_R)) \qquad \forall h\in\mathbb{N},\ R>0. \tag{6.5}$$

Let $\psi_h(x)(t)$ be the unique solutions in $[0,T]$ of the ODE $\dot\psi(t) = b_h(t,\psi(t))$ with the initial condition $\psi(0)=x$ and assume that for any $R>0$ there exists a constant C_R such that

$$\left|\int_{B_R} \varphi\left(\psi_h(x)(t)\right)\,dx\right| \le C_R \int_{B_{M_R}} |\varphi(y)|\,dy \qquad \forall\varphi\in C_c(\mathbb{R}^d),\ t\in[0,T] \tag{6.6}$$

with $M_R = R + T\sup_h\|b_h\|_\infty$. Then, denoting by ψ the regular Lagrangian flow relative to b starting from $\mathbb{R}^d$ given by Theorem 6.2, the functions ψ_h converge in $L^1_{\rm loc}(\mathbb{R}^d;\mathscr{S}(\mathbb{R}^d))$ to the function ψ, i.e.

$$\lim_{h\to\infty}\int_{B_R}\sup_{[0,T]}|\psi(x)-\psi_h(x)|\,dx = 0 \qquad \forall R>0.$$

Proof. We fix $R>0$ and notice that, by the uniform bound on $\|b_h\|_\infty$, the set $\cup_h\psi_h(B_R)$ is relatively compact in $\mathscr{S}(\mathbb{R}^d)$. Theorem 6.5 ensures the existence of a subsequence (not relabelled) such that ψ_h converge to some Young measure γ_x in $\mathscr{S}(\mathbb{R}^d)$, i.e.

$$\lim_{h\to\infty}\int_{B_R}\phi(x,\psi_h(x))\,dx = \int_{B_R}\int_{\mathscr{S}(\mathbb{R}^d)}\phi(x,\psi)\,d\gamma_x(\psi)\,dx$$

for any bounded function $\phi(x,\psi) : B_R\times\mathscr{S}(\mathbb{R}^d)\to\mathbb{R}$ continuous with respect to ψ and $\mathscr{L}^d$-measurable with respect to x. Using functions $\phi(x,\psi)=\varphi(\psi(t))$, with $\varphi\in C_c(\mathbb{R}^d)$ and $t\in[0,T]$, we can pass to the limit as $h\to\infty$ in (6.6) to obtain

$$\left|\int_{B_R}\int_{\mathscr{S}(\mathbb{R}^d)}\varphi\left(\psi(t)\right)\,d\gamma_x(\psi)\,dx\right| \le C_R\int_{B_{M_R}}|\varphi(y)|\,dy \qquad \forall\varphi\in C_c(\mathbb{R}^d),\ t\in[0,T]. \tag{6.7}$$

This inequality, with the same argument leading to (6.2), gives

$$\int_0^T\int_{\mathscr{S}(\mathbb{R}^d)}\chi_{N_t}\circ\psi(t)\,d\gamma_x(\psi)\,dt = 0 \qquad \text{for } \mathscr{L}^d\text{-a.e. } x\in B_R \tag{6.8}$$

for any $\mathscr{L}^{d+1}$-negligible Borel set $N\subset(0,T)\times\mathbb{R}^d$, with $N_t=\{x:\ (t,x)\in N\}$.

We claim now that $\gamma_x\left(\mathscr{S}(\mathbb{R}^d)\setminus\mathscr{S}^b_x(\mathbb{R}^d)\right)=0$ for $\mathscr{L}^d$-a.e. $x\in B_R$. Indeed, by a continuity argument it suffices to show that for any $t\in[0,T]\cap\mathbb{Q}$ we have

$$\psi(t) = x + \int_0^t b(\tau,\psi(\tau))\,d\tau \qquad \text{for } \gamma_x\text{-a.e. } \psi\in\mathscr{S}(\mathbb{R}^d)$$

for $\mathscr{L}^d$-a.e. $x \in B_R$. In order to show this fact we fix $t \in [0,T]$, a smooth and bounded vector field c and define

$$\phi(x,\psi) := \left|\psi(t) - x - \int_0^t c(\tau,\psi(\tau))\,d\tau\right|$$

to obtain

$$\begin{aligned}
&\int_{B_R}\int_{\mathscr{S}(\mathbb{R}^d)}\left|\psi(t) - x - \int_0^t c(\tau,\psi(\tau))\,d\tau\right|\,d\gamma_x(\psi)\,dx \\
=\;& \lim_{h\to\infty}\int_{B_R}\left|\int_0^t (b_h - c)(\tau,\psi_h(x)(\tau))\,d\tau\right|\,dx \\
\le\;& \limsup_{h\to\infty}\int_0^t\int_{B_R}|b_h - c|(\tau,\psi_h(x)(\tau))\,dx\,d\tau \\
\le\;& C_R\limsup_{h\to\infty}\int_{(0,T)\times B_{M_R}}|b_h(\tau,y) - c(\tau,y)|\,d\tau\,dy = C_R\|b-c\|_{L^1((0,T)\times B_{M_R})}.
\end{aligned}$$

Choosing $c = c_\epsilon = b * \rho_\epsilon$ and using Fatou's Lemma we obtain

$$\int_{B_R}\liminf_{\epsilon\downarrow 0}\int_{\mathscr{S}(\mathbb{R}^d)}\left|\psi(t) - x - \int_0^t c_\epsilon(\tau,\psi(\tau))\,d\tau\right|\,d\gamma_x(\psi)\,dx = 0.$$

Let $N = S_b \cup \left\{(x,t) \notin S_b :\ b(t,x) \neq \tilde{b}(t,x)\right\}$, let $N_\tau = \{x :\ (\tau,x) \in N\}$ and fix $x \in B_R$ where the lim inf inside the integral is 0 and $\gamma_x\left(\{\psi : \int_0^T \chi_{N_\tau}\circ\psi(\tau)\,d\tau > 0\}\right) = 0$, and choose a sequence $(\epsilon_k) \downarrow 0$ such that

$$\lim_{k\to\infty}\int_{\mathscr{S}(\mathbb{R}^d)}\left|\psi(t) - x - \int_0^t c_{\epsilon_k}(\tau,\psi(\tau))\,d\tau\right|\,d\gamma_x(\psi) = 0.$$

Using the pointwise convergence of c_{ϵ_k} to $\tilde{b} = b$ out of N we obtain

$$\int_{\mathscr{S}(\mathbb{R}^d)}\left|\psi(t) - x - \int_0^t b(\tau,\psi(\tau))\,d\tau\right|\,d\gamma_x(\psi) = 0.$$

Thanks to the previous claim and to (6.7) we can apply Theorem 5.4 to obtain that $\gamma_x = \delta_{\psi(x)}$ for $\mathscr{L}^d$-a.e. $x \in B_R$. From (6.7) we obtain also that ψ is the regular Lagrangian flow starting from B_R. By Remark 6.3, which ensures the possibility of finding an approximation $b_h = b * \rho_{\epsilon_h}$ of b for which (6.6) holds, this proves also Theorem 6.2. Moreover, if (5.6) holds then the injectivity of the flow is shown in Theorem 5.7, in the more general context of flows induced by probability measures γ_x in $\mathscr{S}(\mathbb{R}^d)$.

In order to show L^1 convergence of ψ_h to ψ on B_R we use the test function $\phi(x,\psi) = \sup_{[0,T]} |\psi - \psi(x)|$ to obtain

$$\lim_{h\to\infty} \int_{B_R} \sup_{[0,T]} |\psi_h(x) - \psi(x)|\, dx = \int_{B_R} \int_{\mathscr{S}(\mathbb{R}^d)} \phi(x,\psi)\, d\gamma_x(\psi)\, dx = 0.$$

□

Since sequences converging in L^1 have subsequences converging a.e. the stability theorem and a diagonal argument provide also a subsequence $(\psi_{h(k)})$ such that

$$\lim_{k\to\infty} \psi_{h(k)}(x) = \psi(x) \qquad \text{uniformly in } [0,T] \text{ for } \mathscr{L}^d\text{-a.e. } x \in \mathbb{R}^d.$$

Remark 6.7 (A comparison with the DiPerna–Lions approach). It is interesting to compare our characterization of Lagrangian flows with the one given in [26] (see in particular Theorem III.1 and the remarks following it, and Theorem III.2), based on the following three properties (here we use the standard notation $\psi(x)(s) = \psi(x,s)$):

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

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

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

(c) for $\mathscr{L}^d$-a.e. x we have $\psi\,(\psi(x,s),t) = \psi(x,s+t)$ for all $s,\, t \ge 0$ with $s+t \le T$.

Notice that our definition of Lagrangian flows (see Definition 6.1) does not involve the semigroup property (c) and weakens (b), by requiring only upper bounds on the flow. Condition (a), although stated in distribution sense, is equivalent to our condition that $\psi(x,\cdot)$ is an integral solution of the ODE for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$ (possibly by modifying ψ in a $\mathscr{L}^{d+1}$-negligible set, see for instance Proposition 3.105 in [6]).

In our setting of regular Lagrangian flows starting from A, condition (c) can be recovered as a consequence with the following argument: fix $s \in \mathbb{Q}\cap[0,T]$, a $\mathscr{L}^d$-measurable set $A_h \subset \mathbb{R}^d$ such that (6.1) holds for some constant C_h and define

$$\tilde\psi(x,t) := \begin{cases} \psi(x,t) & \text{if } t \le s; \\ \\ \psi\,(\psi(x,s),t-s) & \text{if } s \le t \le T. \end{cases}$$

It is immediate to check that $\tilde\psi(x,\cdot)$ is an integral solution of the ODE for $\mathscr{L}^d$-a.e. $x \in A_h$ and that

$$\int_{A_h} \varphi\left(\tilde\psi(x,t)\right)\, dx \le C_h^2 \int_{\mathbb{R}^d} \varphi(y)\, dy \qquad \forall \varphi \in C_c(\mathbb{R}^d)$$

for any $t \geq s$. Then, Theorem 6.4 gives $\tilde{\psi}(x,\tau) = \psi(x,\tau)$ for any $\tau \in [0,T]$, for $\mathscr{L}^d$-a.e. $x \in A_h$. Since s is arbitrary, a density argument gives

$$\psi\left(\psi(x,s),t\right) = \psi(x,s+t) \quad \text{for all } s,\, t \geq 0 \text{ with } s+t \leq T, \text{ for } \mathscr{L}^d\text{-a.e. } x \in A_h.$$

Finally, we use the fact (by definition) we can cover $\mathscr{L}^d$-almost all of A by a sequence of sets A_h with the properties above.

Acknowledgements. I wish to thank the University of Nice, where part of this work was done, and in particular L.Almeida and Y.Brenier for their invitation. I also thank C.Mantegazza for some comments on a preliminary version of this paper and one of the referees of the paper for his helpful comments.

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] G.Alberti: Personal communication.

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

[7] L.Ambrosio, B.Kirchheim, M.Lecumberry & T.Riviere: *On the rectifiability of defect measures arising in a micromagnetics model.* Nonlinear Problems in Mathematical Physics and related topics II, in honor of O.A. Ladyzhenskaya. M.S. Birman, S. Hildebrandt, V.A. Solonnikov and N.Uraltseva Eds., International Mathematical Series, Kluwer/Plenum, 2002 (pages 29–60).

[8] 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.

[9] L.Ambrosio, F.Bouchut & C.De Lellis: *Well-posedness for a class of hyperbolic systems of conservation laws in several space dimensions.* Preprint, 2003 (submitted to Comm. PDE and available at http://cvgmt.sns.it).

[10] 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.

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

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

[13] A.Bressan: *An ill posed Cauchy problem for a hyperbolic system in two space dimensions.* Preprint, 2003.

[14] 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.

[15] C.Castaing & M.Valadier: *Convex analysis and measurable multifunctions.* Lecture Notes in Mathematics **580**, Springer, 1977.

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

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

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

[19] F.Colombini & N.Lerner: *Uniqueness of L^∞ solutions for a class of conormal BV vector fields.* Preprint, 2003.

[20] F.Colombini & J.Rauch: *Unicity and nonunicity for nonsmooth divergence free transport.* Preprint, 2003.

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

[22] M.Cullen & M.Feldman: *Lagrangian solutions of semigeostrophic equations in physical space.* To appear.

[23] C.De Lellis, F.Otto & M.Westdickenberg: *Structure of entropy solutions to the eikonal equation.* To appear on the Journal of the European Mathematical Society.

[24] C.De Lellis & F.Otto: *Structure of entropy solutions for multi-dimensional conservation laws.* To appear on Archive for Rational Mechanics & Analysis.

[25] N.De Pauw: *Non-unicité du transport par un champ de vecteurs presque BV*. Available at http://math.polytechnique.fr/seminaires/seminaires-edp/2002-2003 .

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

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

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

[29] M.Hauray: *On Liouville transport equation with potential in $BV_{\rm loc}$*. (2003) To appear on Comm. in PDE.

[30] M.Hauray: *On two-dimensional Hamiltonian transport equations with $L^p_{\rm loc}$ coefficients.* (2003) To appear on Ann. Nonlinear Analysis IHP.

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

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

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

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