LIPSCHITZ REGULARITY AND APPROXIMATE DIFFERENTIABILITY OF THE DIPERNA-LIONS FLOW

LUIGI AMBROSIO, MYRIAM LECUMBERRY, AND STEFANIA MANIGLIA

1. Introduction

In a recent paper [7] Le Bris and Lions studied, among other things, the differentiability properties of the flow $X(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ associated to a vectorfield $b : (0,T) \times \mathbb{R}^d \to \mathbb{R}^d$ having a Sobolev regularity with respect to the space variable. Under suitable global conditions on b analogous to those considered in [8], where the flow X has been first characterized, they show that

$$\frac{X(t,x+\varepsilon y) - X(t,x)}{\varepsilon} \quad \longrightarrow \quad Z(t,x,y) \tag{1.1}$$

where the convergence, as $\varepsilon \downarrow 0$, occurs with respect to the local convergence in measure in $\mathbb{R}^d_x \times \mathbb{R}^d_y$, and uniformly in time. The map $Z(t,x,y)$ can be considered, according to this limiting procedure, a kind of "derivative" of the flow $X(t,\cdot)$ at x along the direction y.

This result raises several questions about the nature of Z and the convergence of the difference quotients: the main one is whether we can infer some kind of Lipschitz property of the flow from this convergence. This is indeed closely related to the problem of passing from the local convergence in measure in $\mathbb{R}^d_x \times \mathbb{R}^d_y$ to the a.e. convergence to 0 as $\varepsilon \downarrow 0$ of the quantities

$$\int_{B_R(0)} 1 \wedge \left| \frac{X(t,x+\varepsilon y) - X(t,x)}{\varepsilon} - Z(t,x,y) \right| dy \qquad R > 0. \tag{1.2}$$

Notice that elementary Fubini-type arguments show that this passage is possible only for a sequence $(\varepsilon_i) \downarrow 0$, but the convergence to 0 of the integrals (1.2) *only* along some sequence (ε_i) does not seem to lead to any kind of Lipschitz property. Indeed, after the completion of this paper, in a joint work of the first author and Malý [4], a new characterization of the convergence (1.1) (for t fixed) is found, involving only the x variable, and an example that no Lipschitz property can be derived from this weak differentiability property is explicitely given.

Assuming for the sake of simplicity in this introductory discussion that b is autonomous and that both b and its divergence are globally bounded, we are able to answer positively these questions under an assumption slightly stronger than $W^{1,1}_{\rm loc}$, namely that the local maximal function of $|\nabla b|$ belongs to $L^1_{\rm loc}$ (this holds if and only if $|\nabla b| \ln(2 + |\nabla b|) \in L^1_{\rm loc}$). Under this assumption we show in Theorem 3.3 that $Z(t,x,y)$ is representable as $L(t,x)y$

for suitable linear maps $L(t,x):\mathbb{R}^d\to\mathbb{R}^d$ (see also [4] and Remark 3.7); moreover, for any ball $B_R(0)$ and any $\delta>0$ we can find a Borel set $A\subset B_R(0)$ such that

$$\mathscr{L}^d\left(B_R(0)\setminus A\right)<\delta \quad\text{and}\quad X(t,\cdot)|_A \text{ is a Lipschitz map for any } t\in[0,T].$$

It turns also out that indeed the map $L(t,x)$ can be characterized $\mathscr{L}^{d+1}$-a.e. in $[0,T]\times A$ as the classical differential, given by Rademacher theorem, of any Lipschitz extension of $X(t,\cdot)|_A$ (see also Section 2.1 for a different characterization in terms of the so-called approximate differential). Furthermore, combining "forward" and "backward" Lipschitz estimates we obtain in Theorem 3.4 also bi-Lipschitz estimates, on large sets depending on time.

The countable Lipschitz property immediately implies that several classical identities (known to be true under the assumptions of the Cauchy-Lipschitz theorem), as the *explicit* formula for the density transported by the flow, are still true in this setting, see Corollary 3.5.

The strategy in [7] is based on the analysis of the dimensional flow in $\mathbb{R}^{2d}$

$$Y^\varepsilon(t,x,y):=\left(X(t,x),\frac{X(t,x+\varepsilon y)-X(t,x)}{\varepsilon}\right)$$

associated to the vector fields

$$\left(b(x),\frac{b(x+\varepsilon y)-b(x)}{\varepsilon}\right)$$

and on the theory of renormalized solutions for the limit vectorfield $(b(x),\nabla b(x)y)$ (see also [10] for related results in a BV context). Our strategy still uses the same difference quotients, but does not require this extension of the theory. Our starting point has been the observation that, in a smooth setting, the time derivative of $\ln|\nabla X(t,\cdot)|$ can be controlled by $|\nabla b|(X(t,x))$; looking for a suitable discrete counterpart of this fact we considered the quantities (here and in the sequel $⨍$ denotes the averaged integral)

$$\tilde\beta_t^\varepsilon(x):=⨍_{B_\varepsilon(x)} f\left(\frac{|X(t,y)-X(t,x)|}{\varepsilon}\right)dy,$$

where $f(s)$ is of the form $\ln(1+s\wedge\lambda)$ for some $\lambda\geq 0$. Their formal limit is

$$⨍_{B_1(0)} f\left(|\nabla X(t,x)y|\right)dy,$$

a quantity comparable to $f(|\nabla X(t,x)|)$. Then we consider the push-forward β_t^ε of $\tilde\beta_t^\varepsilon$ under the map $X(t,\cdot)$ and the push forward $w_t^\varepsilon(x,y)$ of $\chi_{B_R(0)}(x)\chi_{B_1(0)}(y)$ under the map $Y^\varepsilon(t,\cdot)$ to obtain that β_t^ε satisfy a transport inequality

$$\frac{d}{dt}\beta_t^\varepsilon+D_x\cdot(b\beta_t^\varepsilon)\leq r_t^\varepsilon \quad\text{with}\quad r_t^\varepsilon(x):=\lambda^d⨍_{B_\lambda(0)}\frac{|b(x+\varepsilon y)-b(x)|}{\varepsilon|y|}w_t^\varepsilon(x,y)\,dy,$$

whose right hand side can be controlled by the maximal function of $|\nabla b|$. Standard representation results for the solutions of transport problems then give estimates from above on β_t^ε and then on $\tilde\beta_t^\varepsilon$.

It is not clear whether our argument can be improved, getting Lipschitz properties in the $W^{1,1}_{\rm loc}$ case, or even in the $BV_{\rm loc}$ case considered in [2]. Some extensions of our result, together with some other open problems, are discussed in Remark 3.8.

2. Notation and preliminary results

Given a map $w(t,x)$ depending on time and space, we will systematically use the notation w_t for the map $x \mapsto w(t,x)$, while a derivative with respect to time will be denoted by $\dot f$ in the case of ODE's and by $\frac{d}{dt}f$ in the case of PDE's. The least Lipschitz constant of a Lipschitz function f will be denoted by $\mathrm{Lip}\, f$.

We denote by $\mathscr{L}^d$ the Lebesgue measure in $\mathbb{R}^d$ and by ω_d the Lebesgue measure of the unit ball of $\mathbb{R}^d$. Recall that a sequence of Borel maps (f_h) is said to be locally convergent in measure to f if

$$\lim_{h\to\infty} \mathscr{L}^d\left(\{x \in B_R(0):\ |f_h(x) - f(x)| > \delta\}\right) = 0 \qquad \forall R > 0,\, \delta > 0.$$

Equivalently, one can say that $1 \wedge |f_h - f| \to 0$ in $L^1_{\rm loc}(\mathbb{R}^d)$.

2.1. Approximate differentiability.

We start by recalling the classical definition of approximate differentiability: a Borel map $X : \mathbb{R}^d \to \mathbb{R}^m$ is said to be *approximately differentiable* at $x \in \mathbb{R}^d$ if there exists a linear map $L : \mathbb{R}^d \to \mathbb{R}^m$ such that the difference quotients

$$y \mapsto \frac{X(x+\varepsilon y) - X(x)}{\varepsilon}$$

locally converge in measure as $\varepsilon \downarrow 0$ to Ly. This is obviously a local property and we still denote by $\nabla X(x)$ the approximate differential whenever no ambiguity arises. The approximate differentiability condition can also be stated in a seemingly stronger but equivalent way, by saying that there is a map $\tilde X$, differentiable in the classical sense at x, such that $\tilde X(x) = X(x)$ and the coincidence set $\{y :\ X(y) = \tilde X(y)\}$ has density 1 at x. The latter formulation can be used, in conjunction with Rademacher theorem, to show that if $X|_A$ is a Lipschitz map for some set $A \subset \mathbb{R}^d$, then X is approximately differentiable at $\mathscr{L}^d$-almost any point of A: it suffices to find a Lipschitz extension $\tilde X$ to the whole of $\mathbb{R}^d$ of $X|_A$ (see for instance 2.10.43 of [9]) to obtain the approximate differentiability property at any point of density 1 of A where $\tilde X$ is classically differentiable. It is worth to mention also (see 3.1.8 of [9]) a converse statement: approximate differentiability at any point of a Borel set A implies that we can cover A by an increasing family of Borel sets A_h such that the restriction of $X|_{A_h}$ is a Lipschitz map for any h.

In connection with Sobolev (or even BV) functions, the following classical result holds (see for instance [1], Lemma 3.81 and Theorem 3.83):

Theorem 2.1 (Approximate differentiability of Sobolev functions)**.** *Let $\Omega \subset \mathbb{R}^d$ be an open set and let $f \in W^{1,1}_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$. Then we have*

$$\lim_{r\downarrow 0} \fint_{B_r(x)} \frac{|f(y)-f(x)-\nabla f(x)(y-x)|}{|y-x|}\,dy = 0 \quad \text{for } \mathscr{L}^d\text{-a.e. } x\in\Omega. \tag{2.1}$$

Furthermore

$$\fint_{B_r(x)} \frac{|f(y)-f(x)|}{|y-x|}\,dy \le \int_0^1 \fint_{B_{tr}(x)} |\nabla f|(y)\,dydt \qquad \text{for any ball } B_r(x)\subset\subset\Omega. \tag{2.2}$$

In the following theorem we state a basic criterion for approximate differentiability: basically it says that if the asymptotic L^1 norm of truncated difference quotients can be bounded independently of the truncation level, then the map is approximately differentiable. More precisely, in order to study the Lipschitz properties of the flow, we are going to apply Remark 2.3 with $f(t) = \ln(1+t\wedge\lambda)$ with λ sufficiently large.

Theorem 2.2. *Let $f_i : [0,+\infty) \to [0,+\infty)$ be subadditive and nondecreasing functions such that $\sup_i \sup f_i = +\infty$, and let $X : \mathbb{R}^d \to \mathbb{R}^m$ be a Borel map. Assume that*

$$\limsup_{i\to\infty} \limsup_{r\downarrow 0} \fint_{B_r(x)} f_i\left(\frac{|X(y)-X(x)|}{r}\right)dy < +\infty \qquad \forall x\in A$$

for some Borel set $A\subset\mathbb{R}^d$. Then X is approximately differentiable at $\mathscr{L}^d$-a.e. $x\in A$.

Proof. We denote by $c(x)$ the double limsup appearing in the statement and we assume with no loss of generality that $\mathscr{L}^d(A) < +\infty$. Since $c(x)$ is finite for any $x\in A$, for any $\varepsilon>0$ we can find a compact set $K\subset A$ and $M\in\mathbb{R}$ such that $\mathscr{L}^d(A\setminus K)<\varepsilon$ and $c\le M-1$ on K, $|X|\le M$ on K. Furthermore, by applying Egorov theorem to the family of functions

$$g_k(x) := \sup_{i\ge k} \limsup_{r\downarrow 0} \fint_{B_r(x)} f_i\left(\frac{|X(y)-X(x)|}{r}\right)dy \qquad x\in K$$

we can find a compact set $K'\subset K$ satisfying $\mathscr{L}^d(K\setminus K')<\varepsilon$ such that

$$\limsup_{r\downarrow 0} \fint_{B_r(x)} f_i\left(\frac{|X(y)-X(x)|}{r}\right)dy < M \qquad \forall x\in K'$$

for i sufficiently large independent of x. Denoting by c_d the Lebesgue measure of the intersection of two open balls with radius 1 whose distance between the centers is 1, we choose i in such a way that

$$f_i(\lambda_M) > \frac{2M\omega_d}{c_d} \text{ for some } \lambda_M \ge 0$$

and we apply in an analogous way Egorov theorem again to find a compact set $K'' \subset K'$ such that $\mathscr{L}^d(K' \setminus K'') < \varepsilon$ and

$$\fint_{B_r(x)} f_i\left(\frac{|X(y)-X(x)|}{r}\right) dy \le M \qquad \forall x \in K'' \tag{2.3}$$

for $r < r_0$, with $r_0 > 0$ independent of x. Notice that by construction $\mathscr{L}^d(A \setminus K'') < 3\varepsilon$.

We now claim that the restriction of X to K'' is a Lipschitz map. Indeed, for any pair of points $x, y \in K''$ we can estimate $|X(x) - X(y)|$ with $2M|x-y|/r_0$ if $|x-y| \ge r_0$. If $r := |x-y| < r_0$ we apply (2.3) twice and the subadditivity of f_i to obtain

$$\frac{1}{\omega_d r^d}\int_{B_r(x)\cap B_r(y)} f_i\left(\frac{|X(x)-X(y)|}{r}\right) dz \le$$
$$\fint_{B_r(x)} f_i\left(\frac{|X(z)-X(y)|}{r}\right) dz + \fint_{B_r(y)} f_i\left(\frac{|X(z)-X(x)|}{r}\right) dz \le 2M.$$

Since $\mathscr{L}^d(B_r(x) \cap B_r(y)) = c_d r^d$ we obtain

$$f_i\left(\frac{|X(x)-X(y)|}{r}\right) \le \frac{2\omega_d M}{c_d},$$

so that our choice of λ_M and the monotonicity of f_i give

$$|X(x)-X(y)| \le \lambda_M |x-y|.$$

□

Remark 2.3. Let $f : [0,+\infty) \to [0,+\infty)$ be a subadditive and nondecreasing function. The argument used in the proof of Theorem 2.1 shows that the conditions

$$\sup_{r\in(0,r_0)} \fint_{B_r(x)} f\left(\frac{|X(y)-X(x)|}{r}\right) dy \le M \quad \text{and} \quad |X| \le M_1 \quad \text{on } A$$

for some $M \ge 0$, $M_1 \ge 0$, $r_0 > 0$ imply that

$$\mathrm{Lip}(X|_A) \le \max\{\frac{2M_1}{r_0}, \lambda\}$$

provided $f(\lambda) > 2M\omega_d/c_d$.

2.2. **Maximal functions.** Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ be a nonnegative function. The *local maximal function* $f^\sharp$ is defined by

$$f^\sharp(x) := \sup_{t\in(0,1)} \fint_{B_t(x)} f(y)\, dy.$$

It is well known (see for instance [12]) that the weak L^1 estimate

$$\mathscr{L}^d\left(\{x \in B_R(0) :\ f^\sharp(x) > \lambda\}\right) \le \frac{C(d)}{\lambda}\int_{B_{R+1}(0)\cap\{f>\lambda\}} f(y)\, dy \qquad \forall \lambda > 0$$

gives that $f^\sharp$ is finite $\mathscr{L}^d$-a.e., and that

$$\int_{B_R(0)} f^{\sharp p}\,dx \le \frac{C(d)p2^{2p}}{p-1}\int_{B_{R+1}(0)} |f|^p\,dx \qquad \forall p\in(1,\infty). \tag{2.4}$$

In the critical case $p=1$ we have

$$\int_{B_R(0)} f^\sharp\,dx \le \omega_d R^d + C(d)\int_{B_{R+1}(0)} f\ln(2+f)\,dx. \tag{2.5}$$

2.3. **Flow associated to a vectorfield.** In this section we consider a vector field $B(t,z)=B_t(z)$ satisfying the following conditions:

[P1] $B\in L^1\left([0,T];W^{1,p}_{\mathrm{loc}}(\mathbb{R}^m;\mathbb{R}^m)\right)$;

[P2] $\frac{|B|}{1+|z|}\in L^1\left([0,T];L^1(\mathbb{R}^m)\right)+L^1\left([0,T];L^\infty(\mathbb{R}^m)\right)$;

[P3] $[\operatorname{div} B_t]^-\in L^1\left([0,T];L^\infty(\mathbb{R}^m)\right)$.

We denote by L the constant

$$L:=e^{\int_0^T \|[\operatorname{div} B_t]^-\|_\infty\,dt}. \tag{2.6}$$

If also

[P4] $[\operatorname{div} B_t]^+\in L^1\left([0,T];L^\infty(\mathbb{R}^m)\right)$

holds, we set

$$\tilde L:=e^{\int_0^T \|[\operatorname{div} B_t]^+\|_\infty\,dt}. \tag{2.7}$$

The following definition of flow is a variant of the one adopted in [8], as it does not involve the semigroup property. Basically in this definition the flow is considered as a measurable map $x\mapsto X(\cdot,x)$ with values in the space of continuous maps, while in [8] it is considered as a continuous map $t\mapsto X(t,\cdot)$, with a suitable metric in the space of measurable maps in $\mathbb{R}^d$ that induces the convergence in measure. See Remark 6.7 of [2] for the proof of the equivalence between the two definitions, at least under the assumptions [P1], [P2], [P3].

Definition 2.4 (Flow)**.** *We say that $Y(t,z):[0,T]\times\mathbb{R}^m\to\mathbb{R}^m$ is a flow (starting at time 0) relative to a vectorfield $B(t,z)$ if the following two conditions are satisfied:*

(a) *for $\mathscr{L}^m$-a.e. $z\in\mathbb{R}^m$ the map $t\mapsto Y(t,z)$ is an absolutely continuous integral solution of the ODE $\dot\gamma=B(t,\gamma)$ in $[0,T]$, with $\gamma(0)=z$.*

(b) *$Y(t,\cdot)_\#\mathscr{L}^m\le C\mathscr{L}^m$ for some constant C independent of t.*

An important consequence of condition (b), that we will use later on, is the property

$$\int_{\mathbb{R}^m}\mathscr{L}^1\left(\{t\in(0,T):\ (t,Y(t,x))\in N\}\right)dx=0 \tag{2.8}$$

for any $\mathscr{L}^{m+1}$-negligible set $N\subset(0,T)\times\mathbb{R}^m$. Indeed, denoting by N_t the t-sections of N, we have that the integral above equals

$$\int_0^T\mathscr{L}^m\left(\{x\in\mathbb{R}^m:\ Y(t,x)\in N_t\}\right)dt=\int_0^T Y(t,\cdot)_\#\mathscr{L}^m(N_t)\,dt=0,$$

where in the last equality we used the fact that N_t is $\mathscr{L}^m$-negligible for $\mathscr{L}^1$-a.e. $t\in(0,T)$.

Theorem 2.5. *Under assumptions [P1], [P2], [P3] there exists a flow, uniquely determined in* $[0,T]\times\mathbb{R}^m$ *up to sets with* $\mathscr{L}^m$*-negligible projection on* $\mathbb{R}^m$*. Moreover, property (b) holds with* $C=L$*, the constant defined in* (2.6)*. The flow has also the following additional properties:*

(a) *there exist vectorfields* B_h*, smooth with respect to the space variable, such that*

$$\|[\operatorname{div} B_{ht}]^{\pm}\|_\infty \le \|[\operatorname{div} B_t]^{\pm}\|_\infty, \tag{2.9}$$

$$\frac{|B_h|}{1+|z|}\in L^1\left([0,T];L^\infty(\mathbb{R}^m)\right),\qquad B_h\in L^1\left([0,T];W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^m;\mathbb{R}^m)\right)$$

and such that the classical flows Y_h *associated to* B_h *satisfy*

$$\lim_{h\to\infty}\int_{B_R(0)}\max_{t\in[0,T]}|Y_h(t,z)-Y(t,z)|\wedge 1\,dz=0\qquad \forall R>0.$$

(b) *If [P4] holds, then* $Y(t,\cdot)_\#\mathscr{L}^m\ge \tilde{L}^{-1}\mathscr{L}^m$*, with* $\tilde{L}$ *as in* (2.7)*.*

Proof. The existence of the flow is proved in [8], together with its uniqueness according to the definition of flow adopted therein. Uniqueness according to Definition 2.4 (a priori a weaker one) is proved in [2] for the case of bounded vectorfields and in [3] in the general case. Statement (a) is proved in [8] (see also [3]) by taking as B_h the standard mollifications of B w.r.t. the space variable. Statement (b) can be easily proved by approximation, using the explicit expression for the densities of $Y_h(t,\cdot)_\#\mathscr{L}^d$, namely

$$\frac{1}{\det\nabla Y_h\left(t,[Y_h(t,\cdot)]^{-1}(x)\right)}.$$

Since $\Delta_h(t,x):=\det\nabla Y_h(t,x)$ solves the ODE $\Delta_h'(t,x)=(\operatorname{div} B_{ht}(Y_h(t,x)))\Delta_h(t,x)$, taking (2.9) into account with the positive parts we obtain an uniform upper bound on Δ_h and therefore a uniform lower bound on the densities. □

Lemma 2.6 (Logarithmic sup estimate)**.** *Let* Y *be a flow relative to* B*. Then*

$$\int_{B_R(0)}\max_{[0,T]}\ln\left(\frac{1+|Y(t,z)|}{1+R}\right)dz\le\|B\|^*\qquad\forall R>0, \tag{2.10}$$

where $\|B\|^*$ *denotes the infimum of all sums*

$$L\|\frac{B^1}{1+|z|}\|_{L^1(L^1)}+\omega_m R^m\|\frac{B^2}{1+|z|}\|_{L^1(L^\infty)},$$

among all decompositions of $|B|/(1+|z|)$ *and* L *is defined in (2.6).*

Proof. Let w_t be the density of $Y(t,\cdot)_\#\chi_{B_R(0)}\mathscr{L}^m$ w.r.t. $\mathscr{L}^m$ and notice that $\|w_t\|_1=\omega_m R^m$ and $\|w_t\|_\infty\le L$, by property (b) of the flow. Using property (a) of the flow we get

$$\begin{aligned}\int_{B_R(0)}\max_{[0,T]}\ln\left(\frac{1+|Y(t,z)|}{1+R}\right)dz &\le \int_{B_R(0)}\int_0^T\frac{|\dot Y(t,z)|}{1+|Y(t,z)|}\,dtdz\\ &= \int_0^T\int_{B_R(0)}\frac{|B_t(Y(t,z))|}{1+|Y(t,z)|}\,dtdz\le\int_0^T\int_{\mathbb{R}^m}\frac{|B_t|w_t}{1+|z|}\,dzdt.\end{aligned}$$

Splitting $|B|/(1+|z|)$ in the sum of a function in $L^1(L^1)$ and a function in $L^1(L^\infty)$ and minimizing among all possible decompositions we obtain (2.10). □

3. Approximate differentiability of the flow

Lemma 3.1. *Assume that $b:(0,T)\times\mathbb{R}^d\to\mathbb{R}^d$ fulfils [P1], [P2], [P3] and let $X(t,x)$ be the flow associated to b. Let $\varepsilon>0$ and let*

$$Y^\varepsilon(t,x,y):=\left(X(t,x),\frac{X(t,x+\varepsilon y)-X(t,x)}{\varepsilon}\right).$$

Then Y_ε is the flow relative to the vector field $B^\varepsilon(t,x,y)=B^\varepsilon_t(x,y)$ in $\mathbb{R}^{2d}$ defined by

$$B^\varepsilon(t,x,y):=\left(b_t(x),\frac{b_t(x+\varepsilon y)-b_t(x)}{\varepsilon}\right). \tag{3.1}$$

In particular condition (b) is fulfilled with $C=L^2$, where

$$L:=e^{\int_0^T\|[\operatorname{div}b_t]^-\|_\infty\,dt}. \tag{3.2}$$

Proof. It is immediate to check that the condition $\dot X(t,x)=b_t(X(t,x))$ $\mathscr{L}^1$-a.e. in $[0,T]$ for $\mathscr{L}^d$-a.e. x implies that $\dot Y^\varepsilon(t,x,y)=B^\varepsilon_t(Y^\varepsilon(t,x,y))$ $\mathscr{L}^1$-a.e. in $[0,T]$ for $\mathscr{L}^{2d}$-a.e. (x,y) (precisely, for $\mathscr{L}^d$-a.e. x, the property holds for any y). In order to check that $Y^\varepsilon(t,\cdot)_\#\mathscr{L}^{2d}\le L^2\mathscr{L}^{2d}$ (with L as in (3.2)) we write the inequality in an integral form

$$\int_{\mathbb{R}^d}\varphi\left(X(t,x),\frac{X(t,x+\varepsilon y)-X(t,x)}{\varepsilon}\right)dx\le L^2\int_{\mathbb{R}^d\times\mathbb{R}^d}\varphi(x,y)\,dxdy$$

for any nonnegative $\varphi\in C_c(\mathbb{R}^d\times\mathbb{R}^d)$ and we notice that the property is trivially true if $b_t\in C^1$ (indeed, in this case the divergence of $B^\varepsilon_t(x,y)$ is $\operatorname{div}b_t(x)+\operatorname{div}b_t(x+\varepsilon y)$). The general case can be immediately achieved using the integral form and the stability property of the flows with respect to approximations by locally Lipschitz vectorfields, ensured by an application of Theorem 2.5(a) to the vectorfield b. □

Lemma 3.2. *Assume that $b:(0,T)\times\mathbb{R}^d\to\mathbb{R}^d$ fulfils [P1], [P2], [P3] and let $X(t,x)$ be the flow associated to b. Let β be a bounded nonnegative function satisfying*

$$\frac{d}{dt}\beta+D_x\cdot(b\beta)=r\in L^1_{\rm loc}\left([0,T)\times\mathbb{R}^d\right) \tag{3.3}$$

in the sense of distributions, and assume that

(i) $t \mapsto \beta_t$ *is* w^**-continuous between* $[0,T)$ *and* $L^\infty(\mathbb{R}^d)$*;*
(ii) $\sup_{[0,T]\times A} |X(\cdot,x)| < +\infty$ *for some Borel set* $A \subset \mathbb{R}^d$*.*

Then we have

$$\sup_{t\in\mathbb{Q}\cap[0,T)} \beta_t(X(t,x)) \leq L\beta_0(x) + L\int_0^T |r|_s(X(s,x))\,ds$$

for $\mathscr{L}^d$*-a.e.* $x \in A$*.*

Proof. We first extend the PDE to negative times (as done immediately before Theorem 4.1 in [2]), setting $b_t = 0$ and $r_t = 0$ for $t < 0$, and $\beta_t = \beta_0$ for $t < 0$. Accordingly, we set $X(t,x) = x$ for $t < 0$. Then we mollify w.r.t. the space variable both sides to obtain smooth functions $\beta_t^\delta = \beta_t * \rho_\delta$ such that

$$\frac{d}{dt}\beta^\delta + D_x\cdot(b\beta^\delta) \leq |r|^\delta \qquad \text{in } (-\infty,T)\times\mathbb{R}^d$$

with $|r|^\delta \to |r|$ in $L^1_{\rm loc}\big((-\infty,T)\times\mathbb{R}^d\big)$ as $\delta\downarrow 0$ (by the commutator estimate in [8]). Finally we mollify again w.r.t. the time variable both sides to obtain smooth functions $\beta^{\delta,\eta} = (\beta^\delta)*\rho_\eta$ such that

$$\frac{d}{dt}\beta^{\delta,\eta} + D_x\cdot(b\beta^{\delta,\eta}) \leq c^{\delta,\eta} \qquad \text{in } (-\infty,T-\eta)\times\mathbb{R}^d$$

with

$$c^{\delta,\eta} := |r|^\delta * \rho_\eta + |(D_x\cdot(b\beta^\delta)) * \rho_\eta - D_x\cdot(b(\beta^\delta * \rho_\eta))|.$$

Notice that the smoothness of β^δ w.r.t. the spatial variables immediately gives, expanding the spatial divergences, that $c^{\delta,\eta} \to |r|^\delta$ in $L^1_{\rm loc}\big((-\infty,T)\times\mathbb{R}^d\big)$ as $\eta\downarrow 0$. The smoothness of $\beta^{\delta,\eta}$ and the fact that it solves the transport inequality (not only in the distributions sense, but also a.e. in $(-\infty,T-\eta)\times\mathbb{R}^d$)

$$\frac{d}{dt}\beta^{\delta,\eta} + b\cdot\nabla\beta^{\delta,\eta} \leq -\mathrm{div}\, b_t\beta^{\delta,\eta} + c^{\delta,\eta}$$

immediately gives

$$\frac{d}{dt}\Big[e^{\int_{-1}^t \mathrm{div}\, b_\tau(X(\tau,x))\,d\tau}\beta_t^{\delta,\eta}(X(t,x))\Big] \leq e^{\int_{-1}^t \mathrm{div}\, b_\tau(X(\tau,x))\,d\tau} c_t^{\delta,\eta}(X(t,x)) \quad \forall t\in(-1,T-\eta)$$

for a.e. $x \in A$. Precisely, this happens for those $x\in A$ such that the transport inequality above fails at $X(t,x)$ for a set of times t with strictly positive 1-dimensional Lebesgue measure; this set is $\mathscr{L}^d$-negligible by (2.8). Hence, if $\eta\in(0,1)$ we get

$$\begin{aligned}\beta_t^{\delta,\eta}(X(t,x)) &\leq L\beta_{-1}^{\delta,\eta}(x) + L\int_{-1}^t c_s^{\delta,\eta}(X(s,x))\,ds \\ &= L\beta_0*\rho_\delta(x) + L\int_{-1}^t c_s^{\delta,\eta}(X(s,x))\,ds.\end{aligned} \tag{3.4}$$

For any $t \in [0,T)$ we can use the w^*-continuity property of $t \mapsto \beta^\varepsilon_t$ (ensuring the strong convergence of $\beta^{\delta,\eta}_t$ to β^δ_t as $\eta \downarrow 0$) to pass to the limit as $\eta \downarrow 0$ in (3.4), using also condition (b) in Definition 2.4, to obtain

$$\beta^\delta_t(X(t,x)) \le L\beta_0 * \rho_\delta(x) + L\int_{-1}^t |r|^\delta_s(X(s,x))\,ds = L\beta_0 * \rho_\delta(x) + L\int_0^t |r|^\delta_s(X(s,x))\,ds$$

for $\mathscr{L}^d$-a.e. $x \in A$. Passing now to the limit as $\delta \downarrow 0$ and using again condition (b) in Definition 2.4, we eventually obtain

$$\beta_t(X(t,x)) \le L\beta_0(x) + L\int_0^t |r|_s(X(s,x))\,ds \qquad \text{for } \mathscr{L}^d\text{-a.e. } x \in A$$

for any $t \in [0,T)$. By letting t vary in the countable set $\mathbb{Q}\cap[0,T)$, we obtain

$$\sup_{t\in\mathbb{Q}\cap[0,T)} \beta_t(X(t,x)) \le L\beta_0(x) + L\int_0^T |r|_s(X(s,x))\,ds \tag{3.5}$$

for $\mathscr{L}^d$-a.e. $x \in A$. □

Theorem 3.3 (Lipschitz estimate)**.** *Assume that b fulfils [P1] for some $p>1$, [P2], [P3], [P4] and let $X(t,x)$ be the flow associated to b. Then, for any ball $B_R(0)$ and any $\delta>0$ we can find a Borel set $A \subset B_R(0)$ such that $\mathscr{L}^d(B_R(0)\setminus A) < \delta$ and the restriction of $X(t,\cdot)$ to A is a Lipschitz map for any $t\in[0,T]$.*
In particular $X(t,\cdot)$ is approximately differentiable $\mathscr{L}^d$-a.e. in $\mathbb{R}^d$ for any $t \in [0,T]$.

Proof. We consider the flow Y^ε and the associated vectorfield B^ε as in Lemma 3.1 and a ball $B_R(0)$. By Lemma 2.6 we obtain

$$\mathscr{L}^d\left(\{x\in B_R(0):\ \max_{t\in[0,T]}|X(t,x)|>M\}\right) \le \left[\ln\left(\frac{1+M}{1+R}\right)\right]^{-1}\|b\|^*$$

for any $M>R$, hence we can find a constant $M_1>R$ and a Borel set $A_1\subset B_R(0)$ such that $\mathscr{L}^d(B_R(0)\setminus A_1)<\delta/2$ and

$$\max_{[0,T]}|X(t,x)| \le M_1 \qquad \forall x\in A_1. \tag{3.6}$$

We define

$$N := \int_{B_1(0)} \ln(1+|y|)\,dy$$

Since

$$\int_{A_1}\int_0^T |\nabla b_s|(X(s,x))\,ds\,dx \le L\int_0^T\int_{B_{M_1}(0)}|\nabla b_s|(y)\,dyds < +\infty,$$

we can find M_2 such that

$$\mathscr{L}^d(A_1\setminus A)<\delta/2 \quad \text{with} \quad A := \left\{x\in A_1:\ \int_0^T|\nabla b_s|(X(s,x))\,ds < M_2\right\}.$$

Eventually we define $M := NL+\omega_d L^2 M_2$ and choose λ sufficiently large, such that $\ln(1+\lambda) > 2M\tilde{L}/c_d$, where

$$\tilde{L} := e^{\int_0^T \|[\operatorname{div} b_t]^+\|_\infty \, dt}.$$

Notice that by construction $\mathscr{L}^d(B_R(0)\setminus A) < \delta$.

Step 1. We fix the initial measure $\bar\mu = \chi_A(x)\chi_{B_1(0)}(y)\mathscr{L}^{2d}$ to obtain, by Lemma 3.1, that $Y^\varepsilon(t,\cdot)_\#\bar\mu \le L^2\mathscr{L}^{2d}$, with L defined in (3.2). We denote by $w_t^\varepsilon(x,y)$ the density of $Y^\varepsilon(t,\cdot)_\#\bar\mu$ w.r.t. $\mathscr{L}^{2d}$ and notice that $\|w^\varepsilon\|_\infty \le L^2$ and that w^ε solve the following Cauchy problem for the continuity equation:

$$\frac{d}{dt}w^\varepsilon + D_{x,y}\cdot(B^\varepsilon w^\varepsilon) = 0, \qquad w^\varepsilon(0,x,y) = \chi_A(x)\chi_{B_1(0)}(y). \tag{3.7}$$

Moreover, for any nonnegative $\varphi \in C_c(\mathbb{R}^d)$ and any $t\in[0,T]$ we have

$$\int_{\mathbb{R}^d}\varphi(x)\int_{\mathbb{R}^d} w_t^\varepsilon(x,y)\,dydx = \int_{A\times B_1(0)}\varphi(X(t,x))\,dxdy \le L\omega_d\int_{\mathbb{R}^d}\varphi(x)\,dx$$

and therefore

$$\int_{\mathbb{R}^d} w_t^\varepsilon(x,y)\,dy \le L\omega_d \quad \text{for } \mathscr{L}^d\text{-a.e. } x, \text{ for any } t\in[0,T]. \tag{3.8}$$

We are going to apply Remark 2.3 with the function $f(t) := \ln(1+t\wedge\lambda)$. To this aim we define

$$\beta_t^\varepsilon(x) := \int_{\mathbb{R}^d} f(|y|)w_t^\varepsilon(x,y)\,dy.$$

Step 2. (estimates on β^ε) Let $\psi\in C_c^\infty(-2,2)$ with $\psi\equiv 1$ on $[-1,1]$ and let $\psi_R(y)=\psi(y/R)$. Using the test function $\varphi(x)f\psi_R(|y|)$, with $\varphi\in C_c^\infty(\mathbb{R}^d)$, in (3.7) gives

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{R}^d}\varphi(x)\int_{\mathbb{R}^d}(f\psi_R)(|y|)w_t^\varepsilon(x,y)\,dydx \qquad (3.9)\\
= &\int_{\mathbb{R}^d}\langle\nabla\varphi(x),b_t(x)\rangle\int_{\mathbb{R}^d}(f\psi_R)(|y|)w_t^\varepsilon(x,y)\,dydx\\
+ &\int_{\mathbb{R}^d}\varphi(x)\int_{B_\lambda(0)}\frac{\langle b_t(x+\varepsilon y)-b_t(x),y\rangle}{\varepsilon|y|(1+|y|)}\psi_R(|y|)w_t^\varepsilon(x,y)\,dydx\\
+ &\int_{\mathbb{R}^d}\varphi(x)\int_{\mathbb{R}^d}\psi_R'(|y|)\frac{\langle b_t(x+\varepsilon y)-b_t(x),\frac{y}{|y|}\rangle}{\varepsilon}f(|y|)w_t^\varepsilon(x,y)\,dydx
\end{aligned}$$

in the distribution sense in $(0,T)$. Using (3.8), the contribution of $b(x)$ in the last integral in (3.9) can be estimated by

$$\frac{\ln(1+\lambda)\|\psi'\|_\infty}{R\varepsilon}\int_{\mathbb{R}^d}|\varphi(x)||b(x)|\int_{\mathbb{R}^d}w_t^\varepsilon(x,y)\,dydx \le \frac{L\omega_d\ln(1+\lambda)\|\psi'\|_\infty}{R\varepsilon}\int_{\mathbb{R}^d}|\varphi(x)||b(x)|\,dx.$$

Using the inequality $1+|x+\varepsilon y|\le C+2\varepsilon R$ for $x\in\operatorname{supp}\varphi$ and $y\in\operatorname{supp}\psi_R$, and writing $b/(1+|z|)$ as $A+A'$ with $|A|\in L^1\big([0,T];L^1(\mathbb{R}^d)\big)$ and $|A'|\in L^1\big([0,T];L^\infty(\mathbb{R}^d)\big)$ we can also estimate the contribution of $b(x+\varepsilon y)$ in the last integral of (3.9) as follows:

$$\frac{(C+2R\varepsilon)\ln(1+\lambda)\|\psi'\|_\infty}{R\varepsilon}\int_{\mathbb{R}^d}|\varphi(x)|\int_{\{|y|\geq R\}}L^2|A_t|(x+\varepsilon y)+\|A'_t\|_\infty w_t^\varepsilon(x,y)\,dy\,dx.$$

Hence, passing to the limit as $R\to\infty$ in (3.9), the dominated convergence theorem gives

$$\begin{aligned}&\frac{d}{dt}\int_{\mathbb{R}^d}\varphi(x)\beta_t^\varepsilon(x)\,dx\\ =\;&\int_{\mathbb{R}^d}\langle\nabla\varphi(x),b_t(x)\rangle\beta_t^\varepsilon(x)\,dx+\int_{\mathbb{R}^d}\varphi(x)\int_{B_\lambda(0)}\frac{\langle b_t(x+\varepsilon y)-b_t(x),y\rangle}{\varepsilon|y|(1+|y|)}w_t^\varepsilon(x,y)\,dydx.\end{aligned}$$

Since φ is arbitrary this proves that

$$\frac{d}{dt}\beta^\varepsilon+D_x\cdot(b\beta^\varepsilon)\leq r^\varepsilon,\qquad \beta^\varepsilon(0,x)=\chi_A(x)\int_{B_1(0)}f(|y|)\,dy \tag{3.10}$$

with

$$r^\varepsilon(t,x):=\int_{B_\lambda(0)}\frac{|\langle b_t(x+\varepsilon y)-b_t(x),y\rangle|}{\varepsilon|y|^2}w_t^\varepsilon(x,y)\,dy\leq L^2\int_{B_\lambda(0)}\frac{|\langle b_t(x+\varepsilon y)-b_t(x),y\rangle|}{\varepsilon|y|^2}\,dy.$$

By (2.2) in Theorem 2.1 we get

$$r^\varepsilon(t,x)\leq L^2\omega_d\lambda^d|\nabla b_t|^\sharp(x)\qquad\text{for}\qquad\lambda\varepsilon\leq 1. \tag{3.11}$$

Notice that $|\nabla b_t|^\sharp\in L^1_{\rm loc}\left([0,T]\times\mathbb{R}^d\right)$ because of the maximal estimate (2.4). Moreover, by (2.1) and (3.8) we infer

$$\limsup_{\varepsilon\downarrow 0}r_t^\varepsilon(x)\leq L\omega_d|\nabla b_t|(x) \tag{3.12}$$

for $\mathscr{L}^{d+1}$-a.e. $(t,x)\in(0,T)\times\mathbb{R}^d$.

Therefore, by applying Lemma 3.2 and using the definition of A and (3.6), Fatou Lemma gives

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0}\sup_{t\in\mathbb{Q}\cap[0,T)}\beta_t^\varepsilon(X(t,x))\;&\leq\;\limsup_{\varepsilon\downarrow 0}L\beta(0,x)+L\int_0^T|r^\varepsilon|_s(X(s,x))\,ds\\ &\leq\; LN+\omega_dL^2\int_0^T|\nabla b_s|(X(s,x))\,ds<M\end{aligned}$$

for $\mathscr{L}^d$-a.e. $x\in A$. Notice that the bound (3.11) and the integrability of $|\nabla b_t|^\sharp$ are used to ensure that Fatou lemma is applicable.

By Egorov theorem, possibly passing to a slightly smaller set A still satisfying $\mathscr{L}^d(B_R(0)\setminus A)<\delta$, we can assume the existence of $\varepsilon_0>0$ such that

$$\sup_{\varepsilon\in(0,\varepsilon_0)}\sup_{t\in\mathbb{Q}\cap[0,T]}\beta_t^\varepsilon(X(t,x))\leq M\qquad\text{for }\mathscr{L}^d\text{-a.e. }x\in A. \tag{3.13}$$

Step 3. (Conclusion) We now claim that, setting

$$\tilde{\beta}_t^\varepsilon(x) := \fint_{B_1(0)} f\left(\frac{|X(t,x+\varepsilon y)-X(t,x)|}{\varepsilon}\right) dy = \fint_{B_\varepsilon(x)} f\left(\frac{|X(t,y)-X(t,x)|}{\varepsilon}\right) dy$$

the inequality

$$\tilde{\beta}_t^\varepsilon(x) \leq \frac{\tilde{L}}{\omega_d}\beta_t^\varepsilon(X(t,x)) \quad \mathscr{L}^d\text{-a.e. in } A, \text{ for any } t\in[0,T] \tag{3.14}$$

holds. Indeed, recalling that w_t^ε is the density of $Y^\varepsilon(t,\cdot)_\#(\chi_A(x)\chi_{B_1(0)}(y)\mathscr{L}^{2d})$ w.r.t. $\mathscr{L}^{2d}$, we have the identity

$$\int_{\mathbb{R}^d}\varphi(x)\beta_t^\varepsilon(x)\,dx = \int_{\mathbb{R}^d\times\mathbb{R}^d}\varphi(x)f(|y|)w_t^\varepsilon(x,y)\,dxdy = \omega_d\int_A \varphi(X(t,x))\tilde{\beta}_t^\varepsilon(x)\,dx,$$

that tells us that $\beta_t^\varepsilon\mathscr{L}^d = X(t,\cdot)_\#(\omega_d\tilde{\beta}_t^\varepsilon\chi_A\mathscr{L}^d)$. From [P4] and Theorem 2.5 we obtain $\tilde{\beta}_t^\varepsilon \leq (\tilde{L}/\omega_d)\beta_t^\varepsilon\circ X(t,\cdot)$ $\mathscr{L}^d$-a.e. in A, for any $t\in[0,T]$.

Summing up, from (3.13) and (3.14) we infer

$$\sup_{\varepsilon\in(0,\varepsilon_0)}\sup_{t\in\mathbb{Q}\cap[0,T)}\fint_{B_\varepsilon(x)} f\left(\frac{|X(t,y)-X(t,x)|}{\varepsilon}\right) dy \leq \frac{\tilde{L}M}{\omega_d}$$

for $\mathscr{L}^d$-a.e. $x\in A$ and we can use the continuity in time of the left hand side to replace the sup on $\mathbb{Q}\cap[0,T)$ by a sup on the whole interval $[0,T]$. Thanks to Remark 2.3 this implies that, denoting by B the Borel subset of A where the inequality above is fulfilled, we have

$$\operatorname{Lip} X(t,\cdot)|_B \leq \max\{2M_1/\varepsilon_0,\lambda\}. \tag{3.15}$$

□

In order to obtain bi-Lipschitz estimate we combine forward and backward Lipschitz estimates and use the semigroup property of the flow, as discussed in [8] and [2].

Theorem 3.4 (bi-Lipschitz estimates). *Assume that b fulfils [P1] for some $p>1$, [P2], [P3], [P4] and let $X(t,x)$ be the flow associated to b. Then for any absolutely continuous probability measure ρ in $\mathbb{R}^d$ and any $\delta>0$, $t\in[0,T]$ we can find a set M_t such that $\rho(\mathbb{R}^d\setminus M_t)<\delta$ and $X(t,\cdot)|_{M_t}$ is a bi-Lipschitz map.*

Proof. Given s, $t\in[0,T]$ we denote by $Y(t,s,x)$ the flow associated to b starting from time s (so that the 1-parameter flow in Definition 2.4 with $d=m$, $B=b$ corresponds to $Y(t,0,x)$): for any given s it is characterized by the conditions

$$Y(s,s,x)=x, \quad \frac{d}{dt}Y(t,s,x)=b\left(t,Y(t,s,x)\right), \quad Y(t,s,\cdot)_\#\mathscr{L}^d\leq C_s\mathscr{L}^d$$

with C independent of t. Arguing as in [8] (see also Remark 6.7 in [2]) one can use the characterization of the 1-parameter flows to obtain the semigroup property

$$Y(t,s,x)=Y\left(t,r,Y(r,s,x)\right) \quad \text{for all } t\in[0,T], \text{ for } \mathscr{L}^d\text{-a.e. } x \tag{3.16}$$

for any s, $r \in [0,T]$. Notice that the $\mathscr{L}^d$-negligible exceptional set N_{rs} a priori depends on r, s.

Given $\delta > 0$ we can find a "forward" set A such that $Y(t,0,\cdot)|_A$ is Lipschitz and $\rho(\mathbb{R}^d \setminus A) < \delta/2$. By reversing the time variable we can find a "backward" set A_t such that

$$Y(t,0,\cdot)_{\#}\rho(\mathbb{R}^d \setminus A_t) < \frac{\delta}{2}$$

and $Y(0,t,\cdot)|_{A_t}$ is Lipschitz with constant λ. Finally we define M_t so that

$$M_t := A \cap Y(t,0,\cdot)^{-1}(A_t) \setminus N_{t0}$$

Since $M_t \subset A$ we need only to show lower bounds on $|X(t,x) - X(t,y)|$. To this aim, notice that for $x \in M_t$ we have

$$x = Y(0,0,x) = Y\left(0,t,Y(t,0,x)\right)$$

because, by definition, $M_t \cap N_{t0} = \emptyset$. Therefore, for x, $y \in M_t$, since both $Y(t,0,x)$ and $Y(t,0,y)$ belong to A_t we obtain

$$\begin{aligned} |x-y| &= |Y(0,0,x) - Y(0,0,y)| = |Y\left(0,t,Y(t,0,x)\right) - Y\left(0,t,Y(t,0,y)\right)| \\ &\leq \lambda |Y(t,0,x) - Y(t,0,x)|. \end{aligned}$$

As a consequence $|X(t,x) - X(t,y)| = |Y(t,0,x) - Y(t,0,y)| \geq |x-y|/\lambda$ for x, $y \in M_t$. □

In the following corollary we give an explicit representation of the density of the (absolutely continuous) measures transported by the flow. This enables to compute also integrals of nonlinear functions of the densities, a computation that would be impossible without an explicit representation of the densities themselves.

Corollary 3.5 (Explicit representation of $X(t,\cdot)_{\#}\rho$)**.** *Under the assumptions of Theorem 3.3, for any absolutely continuous probability measure $\rho = f\mathscr{L}^d$ in $\mathbb{R}^d$ and any $t \in [0,T]$ we have that the density w_t of $X(t,\cdot)_{\#}\rho$ w.r.t. $\mathscr{L}^d$ is representable as*

$$w_t = \frac{f}{|\det \nabla X(t,x)|} \circ \left[X(t,\cdot)|_{\Sigma_t}\right]^{-1} \tag{3.17}$$

for a suitable Borel set $\Sigma_t \subset \mathbb{R}^d$ whose complement is $\mathscr{L}^d$-negligible. Furthermore, for any nonnegative Borel functions φ, ψ, we have the change of variables formula

$$\int_{\mathbb{R}^d} \varphi(w_t)\psi\, dy = \int_{\mathbb{R}^d} |\det \nabla X(t,x)| \varphi\left(\frac{f}{|\det \nabla X(t,x)|}\right) \psi(X(t,x))\, dx. \tag{3.18}$$

Proof. We recall the *area formula* for Lipschitz maps (see for instance 3.2.3 of [9]): if $w : A \subset \mathbb{R}^d \to \mathbb{R}^d$ is a Lipschitz map, then

$$\int_A h(x) |\det \nabla w|\, dx = \int_{\mathbb{R}^d} \sum_{x \in A \cap w^{-1}(y)} h(x)\, dy \tag{3.19}$$

for any nonnegative Borel function h. By the remarks made in Section 2.1, this formula still holds for maps that are approximately differentiable at any point of A, since we can cover A by a sequence of sets A_h such that $w|_{A_h}$ is Lipschitz for any h.

Hence, denoting by Σ_1 the set of points where $X(t,\cdot)$ is approximately differentiable, we can apply (3.19) to any Borel set $A \subset \Sigma_1$ with $w = X(t,\cdot)$. By applying the semigroup property (3.16) we obtain a Borel set Σ_2 such that $\mathscr{L}^d(\mathbb{R}^d \setminus \Sigma_2) = 0$ and (with the notation of Theorem 3.4)

$$x = Y(0,0,x) = Y\left(0,t,Y(t,0,x\right) = Y\left(0,t,X(t,x)\right) \qquad \forall x \in \Sigma_2,$$

so that $X(t,\cdot)$ is one to one on Σ_2. Setting $\Sigma = \Sigma_1 \cap \Sigma_2$ we can apply (3.19) with $A = \Sigma_1 \cap X(t,\cdot)^{-1}(E)$ and

$$h = \frac{f\chi_\Sigma}{|\det \nabla X(t,\cdot)|}$$

for any Borel set $E \subset \mathbb{R}^d$ to obtain

$$\int_{X(t,\cdot)^{-1}(E)} f(x)\,dx = \int_E \frac{f}{|\det \nabla X(t,x)|} \circ \left[X(t,\cdot)|_\Sigma\right]^{-1}(y)\,dy.$$

Since E is arbitrary, this proves (3.17). To prove (3.18) we just apply (3.19) again with

$$h(x) = \chi_\Sigma(x)\psi(X(t,x))\varphi\left(\frac{f(x)}{|\det \nabla X(t,x)|}\right).$$

□

In the following theorem we discuss the relation between the approximate differential $\nabla X(t,x)y$ and the derivative $Z(t,x,y)$ of the flow considered in [7].

Theorem 3.6. *Assume that b fulfils [P1] for some $p > 1$, [P2], [P3], [P4], let $X(t,x)$ be the flow associated to b. Then for any $t \in [0,T]$ the difference quotients*

$$Z^\varepsilon(t,x,y) := \frac{X(t,x+\varepsilon y) - X(t,x)}{\varepsilon}$$

locally converge in measure in $\mathbb{R}^d_x \times \mathbb{R}^d_y$ as $\varepsilon \downarrow 0$ to $Z(t,x,y) = \nabla X(t,x)y$.

Proof. By the very definition of approximate differential the vector fields $Z^\varepsilon(t,x,y)$ locally converge in measure in $\mathbb{R}^d_y$ as $\varepsilon \downarrow 0$ to $\nabla X(t,x)y$ for any (t,x) where $\nabla X(t,x)$ is defined. Therefore, since $\nabla X(t,x)$ exists for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$, one more integration w.r.t. x gives the result. □

Remark 3.7. Using the theory of renormalized solutions for vectorfields of the form $B(x,y) = (b(x), \nabla b(x)y)$, developed in [7], one can also show that

$$(X(t,x), \nabla X(t,x)y)$$

is the *unique flow* associated to B, where "flow" is understood in a slightly weaker sense than the one adopted in this paper (due to the fact that the vectorfield B fails in this case to satisfy condition [P3]). Furthermore, even in the $W^{1,1}_{\text{loc}}$ case not covered by our results, one can show using the stability results of [7] that that the component $Z(t,x,y)$ of the flow is still representable as $L(t,x)y$ for suitable linear maps $L(t,x) : \mathbb{R}^d \to \mathbb{R}^d$ (precisely $L(t,x)y$ is the limit in measure of $\nabla X_h(t,x)y$, where X_h are the approximating flows).

Remark 3.8 (Extensions and open problems)**.** (1) As the proof of Theorem 3.3 clearly shows, the $W^{1,p}_{\rm loc}$ regularity for some $p>1$ can be weakened by requiring [P1] with $p=1$, [P2], [P3] and

$$\int_0^T\int_{B_R(0)}|\nabla b_t|^\sharp(x)\,dxdt<+\infty\qquad\forall R>0. \tag{3.20}$$

Equivalently, we may require that

$$\int_0^T\int_{B_R(0)}|\nabla b_t(x)|\ln(2+|\nabla b_t(x)|)\,dxdt<+\infty\qquad\forall R>0.$$

One can also notice, in the same spirit of [5], that the expression of r^ε in (3.11) involves only the *symmetric* difference quotients of b_t, namely those of the form

$$\frac{\langle b_t(x+\varepsilon y)-b_t(x),y\rangle}{\varepsilon|y|}$$

that can be controlled using only the symmetric part of the derivative. Therefore, still keeping [P2] and [P3], [P1] and (3.20) can be replaced by

$$\int_0^T\int_{B_R(0)}|(\nabla b_t)+(\nabla b_t)^T|^\sharp(x)\,dxdt<+\infty\qquad\forall R>0,$$

requiring only that the symmetric part of the distributional derivative of b_t is in $L^1_{\rm loc}$; by the results in [5] the flow is well defined also under these weaker conditions.
(2) The local integrability of the maximal function of b_t plays an essential role in the pointwise estimate of r^ε, necessary in order to apply Lemma 3.2. Therefore, as we said in the introduction, it is not clear whether our results can be extended to the $W^{1,1}$ case or even to the BV case considered in [2].
(3) The argument used in the proof of Theorem 3.3 does not lead to an explicit bound of the Lipschitz constant of $X(t,\cdot)$ as a function of δ. More precisely, assuming for simplicity that $|b|$ is globally bounded, we have clearly that the constant M_1 depends only on R, while M_2 and therefore M can be estimated from above with $C(R)/\delta$. Hence $\lambda\sim e^{C(R)/\delta}$ can be estimated explictly and gives a bound on the L^∞ norm of $|\nabla X|$ on $[0,T]\times A$. On the other hand, due to the application of Egorov theorem, the *global* Lipschitz constant of $X(t,\cdot)|_A$ depends also on ε_0, as (3.15) shows. More precisely, we have

$$|X(t,x)-X(t,y)|\le\lambda|x-y|\qquad\forall x,\,y\in A\text{ with }|x-y|\le\varepsilon_0,\,t\in[0,T].$$

References

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

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

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

[4] L.Ambrosio & J.Malý: *Very weak notions of differentiability.* Preprint, 2005 (available at http://cvgmt.sns.it).
[5] 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.
[6] F.Colombini & N.Lerner: *Uniqueness of continuous solutions for BV vector fields.* Duke Math. J., **111** (2002), 357–384.
[7] 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.
[8] R.J.Di Perna & P.L.Lions: *Ordinary differential equations, transport theory and Sobolev spaces.* Invent. Math., **98** (1989), 511–547.
[9] H.Federer: *Geometric Measure Theory.* Springer, 1969.
[10] N.Lerner: *Transport equations with partially BV velocities.* Preprint, 2004.
[11] P.L.Lions: *Sur les équations différentielles ordinaires et les équations de transport.* C. R. Acad. Sci. Paris Sér. I, **326** (1998), 833–838.
[12] E.M.Stein: *Singular integrals and differentiability properties of functions.* Princeton University Press, 1970.

(Luigi Ambrosio) Scuola Normale Superiore, Piazza dei Cavalieri, 56126 Pisa, Italy
E-mail address: `l.ambrosio@sns.it`

(Myriam Lecumberry) Max-Planck Institut für Mathematik, Inselstrasse 22–26, D-04103 Leipzig, Germany
E-mail address: `lecumberry@mis.mpg.de`

(Stefania Maniglia) Dipartimento di Matematica, Università di Pisa, 56126 Pisa, Italy
E-mail address: `maniglia@mail.dm.unipi.it`