Lecture notes on transport equation and Cauchy problem for BV vector fields and applications

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

1 Introduction

PROPRIETA' DI SEMIGRUPPO

In these notes I would like to describe informally the main results obtained in [6], together with some recent improvements. In that paper I study the well posedness of the continuity equation and of the ODE for vector fields $b(t,x)$ having a low regularity with respect to the spatial variables, precisely a BV (bounded variation) regularity. These results extend the DiPerna–Lions theory, in which a Sobolev regularity was considered. The extension to BV vector fields is crucial in view of the application, described in the last paragraph, to a particular system of conservation law considered by Bressan in [13] and then in two papers of mine [7], [8], the first one written with De Lellis and the second one with Bouchut and De Lellis.

The problem can be presented from two different viewpoints, an Eulerian one and a Lagrangian one one. We will see that there is indeed a close link between the two viewpoints.

Let $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a Borel vector field. The Eulerian problem is the well-posedness of

$$\text{(PDE)} \qquad \dot\mu_t + D_x \cdot (b\mu_t) = 0, \qquad \mu_0 = \mathscr{L}^d \llcorner A, \qquad t \in [0,T]$$

(here $\dot\mu_t$ stands for time derivative), where μ_t is a suitable time-depending family of measures, possibly signed. In all situations that we will consider there will be bounds on b and/or μ_t ensuring that

$$\int_I \left(|\mu_t|(B_R) + \int_{B_R} |b|\, d|\mu_t| \right) dt \qquad \forall I \subset\subset (0,T),\ R > 0,$$

so that the PDE will make sense in the sense of distributions in $(0,T)\times\mathbb{R}^d$.

The Lagrangian problem is the uniqueness of

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

In some situations one might hope for a "generic" uniqueness of the solutions of ODE, i.e. for "almost every" initial datum x. An even weaker requirement is the research of a "selection principle", i.e. a strategy to select for almost every x a solution $\gamma(\cdot,x)$ in such a way that this selection is stable w.r.t. smooth approximations of b. In other words, we would like to know that, whenever we approximate b by smooth vector fields b_h, the classical trajectiories associated to b_h satisfy

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

In these notes I will consider only the forward problem (this allows to state sharper one-sided conditions on the divergence of the vector fields involved) and a bounded time interval $[0,T]$. With the simple necessary adaptations more general time intervals could be considered as well.

Finally, I conclude this short introduction with some bibliographical remarks, that do not pretend to be exhaustive. The first paper where the connection between the Lagrangian and the Eulerian viewpoint is investigated in detail, in a weak setting, is due to DiPerna–Lions [24], that consider a Sobolev dependence (precisely proving uniqueness of L^q solutions of the PDE for $W^{1,p}$ vectorfields, $1\le p\le\infty$), and in [29] Lions extended these results to the piecewise Sobolev case. In [14] Capuzzo Dolcetta and Perthame proved, among other things, that the DiPerna–Lions theiry still works assuming only that the simmetric part of the distributional derivative is in $L^1_{\rm loc}$ (no assumption on the antisymmetric part, that could be only a distribution). In this connection it is also worth to mention the papers [15], [16] by Cellina and Cellina–Vornicescu, that study the differential inclusion $\dot\gamma(t)\in A\,(\gamma(t))$, with A maximal monotone. In this case uniqueness of the ODE holds for $\mathscr{L}^d$-a.e. initial datum x.

A fundamental paper is due to Bouchut [12], where the problem is solved for II order equation $\gamma''(t) = b\,(t,\gamma(t))$ and, more generally, for some equations of Hamiltonian type. In the paper [17] (see also [18]) Colombini–Lerner consider a particular class of BV vector fields, that they call co-normal BV vector fields, that reduce to vector fields analogous to Bouchut's ones after a (local) bi-Lipschitz change of coordinates. Finally, in [6] I got the general case, imposing only BV regularity and absolute continuity of the divergence with respect to $\mathscr{L}^d$.

The literature in this area is very wide, due to the fundamental character of the continuity equation, and the enclosed bibliography is far from being exhaustive. In any case I wish to mention also the papers [27], [28], containing other uniqueness results in the

2-dimensional case, and the papers [31], [11], where generalized characteristics and one-sided Lipschitz conditions on the vector field are taken into account (for characteristics in the non-linear setting of conservation laws, see also [22]).

I believe that we are still quite far from a good understanding of the optimal conditions on the vector field b, but there are some counterexamples, that for the sake of brevity I will not discuss, showing which phenomena can prevent the uniqueness: besides the two counterexamples in [24], I recall also the papers [1], [19], [23].

2 The classical setup: $b \in L^1\left([0,T]; W^{1,\infty}(\mathbb{R}^d)\right)$

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

$$\frac{d}{dt}|X(t,x)-X(t,y)|^2 \le 2\|\nabla b_t\|_\infty |X(t,x)-X(t,y)|^2,$$

so that Gronwall lemma immediately gives

$$\operatorname{Lip}\left(X(t,\cdot)\right) \le \exp\left(\int_0^t \|\nabla b_s\|_\infty \, ds\right). \tag{2.1}$$

Turning to the solutions of the ODE, uniqueness will be proved in a more general setting for positive measure-valued solutions (via the superposition principle) and for signed solutions (via the theory of renormalized solutions), so here we focus on the existence and the representation issue. The representation formula is indeed very simple

$$\mu_t := X(t,\cdot)_\# \bar\mu \tag{2.2}$$

and we need only to check it. Notice first that we need only to check the distributional identity on test functions of the form $\psi(t)\varphi(x)$, so that

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

This means that we have to check that $t \mapsto \langle \mu_t, \varphi\rangle$ belongs to $W^{1,1}(0,T)$ for any $\varphi \in C^\infty_c(\mathbb{R}^d)$ and that its distributional derivative is $\int_{\mathbb{R}^d} \langle b_t, \nabla\varphi\rangle \, d\mu_t$.

We show first that the map is absolutely continuous, and in particular $W^{1,1}(0,T)$; then one needs only to compute the pointwise derivative. For every choice of finitely many pairwise disjoint intervals $(a_i,b_i) \subset [0,T]$ we have

$$\sum_{i=1}^n |X(b_i,x)-X(a_i,x)| \le \int_{\cup_i (a_i,b_i)} |\dot X(t,\gamma)| \, dt \le \int_{\cup_i (a_i,b_i)} \|\nabla b_t\|_\infty \, dt$$

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

$$\sum_{i=1}^{n} |\langle \mu_{b_i} - \mu_{a_i}, \varphi\rangle| \leq \int_{\cup_i (a_i,b_i)} \|\nabla b_t\| \nabla b_t\|_\infty \, dt.$$

The absolute continuity of the integral shows that the right hand side can be made small when $\sum_i (b_i - a_i)$ is small. This proves the absolute continuity. For any x the identity $\dot X(t,x) = b(t, X(t,x))$ is fulfilled for $\mathscr{L}^1$-a.e. $t \in [0,T]$. Then, by Fubini's theorem, we know also that for $\mathscr{L}^1$-a.e. $t \in [0,T]$ the previous identity holds for μ-a.e. x, and therefore the chain rule (8.1) gives

$$\frac{d}{dt}\langle \mu_t, \varphi\rangle = \int_{\mathbb{R}^d} \langle \nabla\varphi(X(t,x)), b(t, X(t,x))\rangle \, d\eta = \langle b_t\mu_t, \nabla\varphi\rangle \quad \mathscr{L}^1\text{-a.e. in } [0,T].$$

In the case when $\bar\mu = \rho\mathscr{L}^d$ we can say something more, proving that the measures μ_t in (2.2) are absolutely continuous w.r.t. $\mathscr{L}^d$ and computing their density. Let us start by recalling the classical *area formula* (see for instance [5]): if $f : \mathbb{R}^d \to \mathbb{R}^d$ is a (locally) Lipschitz map, then

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

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

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

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

$$f_\#(\rho\mathscr{L}^d) = \frac{\rho}{|Jf|} \circ f^{-1}\mathscr{L}^d. \tag{2.3}$$

In our case $f(x) = X(t,x)$ is surely 1-1, onto and Lipschitz by (2.1). It remains to show that $JX(t,\cdot)$ does not vanish: in fact, one can show that

$$\max\left\{ JX(t,x), \frac{1}{JX(t,x)} \right\} \leq \int_0^t \|\mathrm{div}\, b_s\|_\infty \, ds \qquad \text{for } \mathscr{L}^d\text{-a.e. } x \tag{2.4}$$

thanks to the following fact, left as an exercise.

Exercise 2.1. Show that for any $t \geq 0$ we have

$$\dot J X(t,x) = \mathrm{div}\, b_t(x, X(t,x)) J(t,x) \qquad \mathscr{L}^d\text{-a.e. in } \mathbb{R}^d.$$

Hint: use the semigroup property to reduce the proof to the case $t = 0$.

The previous exercise and Fubini's theorem give that $JX(\cdot, x)$ solves a linear ODE with the initial condition $JX(0, x) = 1$ for $\mathscr{L}^d$-a.e. x, whence (2.4) follows.

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

Remark 2.2 (A local variant). The theory outlined above still works under the assumptions

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

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

Exercise 2.3. (a) Show that one can test the PDE with test functions $\varphi \in C^1\left((0,T) \times \mathbb{R}^d\right)$ that are bounded, have a bounded gradient, and whose support has a compact projection in $(0,T)$ (hint: truncate in the space variable).

(b) Show that if b is bounded then any solution μ_t of the PDE has a narrowly continuous representative, i.e. there exists a narrowly continuous family of measures $t \mapsto \tilde{\mu}_t$ such that $\mu_t = \tilde{\mu}_t$ for $\mathscr{L}^1$-a.e. $t \in [0,T]$.

3 ODE uniqueness versus PDE uniqueness

In this section we illustrate some general principles, whose concrete application may depend on specific assumptions on b, relating the uniqueness of the ODE to the uniqueness of the PDE. The first very general criterion is the following.

Theorem 3.1. *Let $A \subset \mathbb{R}^d$ be a Borel set. The following two properties are equivalent:*

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

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

Proof. It is clear that (b) implies (a), just choosing $\bar{\mu} = \delta_x$ and noticing that two different solutions $X(t)$, $\tilde{X}(t)$ of the ODE induce two different solutions of the PDE, namely $\delta_{X(t)}$ and $\delta_{\tilde{X}(t)}$.

The converse implication is less obvious and requires the superposition principle that we are going to describe below: any positive solution of the PDE is always a superposition of solutions of the ODE and therefore, when the latter are unique, also the solutions of the PDE are unique. □

We will use the shorter notation Γ_T for the space $C\left([0,T]; \mathbb{R}^d\right)$.

Definition 3.2 (Superposition solutions). *Let $\eta \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ be a measure concentrated on the set of pairs (x,γ) such that γ is an absolutely continuous solution of the ODE with $\gamma(0)=x$. We define*

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

By a monotone class argument the identity defining μ_t^η holds for any Borel function φ such that $\gamma \mapsto \varphi(\gamma(t))$ is η-integrable (or equivalently any μ_t^η-integrable function φ).

Under the integrability condition

$$\int_0^T \int_{\mathbb{R}^d\times\Gamma_T} |b_t|\, d\eta\, dt < +\infty \tag{3.1}$$

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

$$\sum_{i=1}^n |\gamma(b_i)-\gamma(a_i)| \le \int_{\cup_i(a_i,b_i)} |b_t(\gamma)|\, dt$$

for η-a.e. (x,γ) and therefore an integration with respect to η gives

$$\sum_{i=1}^n |\langle \mu_{b_i}^\eta - \mu_{a_i}^\eta, \varphi\rangle| \le \int_{\cup_i(a_i,b_i)} \int_{\mathbb{R}^d\times\Gamma_T} |b_t|\, d\eta\, dt.$$

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

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

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

Remark 3.3. Actually the formula defining μ_t^η does not contain x, and so it involves only the projection of η on Γ_T. Therefore one could also consider measures σ in Γ_T, concentrated on the set of solutions of the ODE. These two viewpoints are basically equivalent: given η one can build σ just by projection, and given σ one can consider the conditional probability measures σ_x induced by the random variable $\gamma \mapsto \gamma(0)$ in Γ_T, the law $\bar\mu$ (i.e. the push forward) of the same random variable and recover η as follows:

$$\int_{\mathbb{R}^d\times\Gamma_T} \varphi(x,\gamma)\, d\eta(x,\gamma) := \int_{\mathbb{R}^d} \left(\int_{\Gamma_T} \varphi(x,\gamma)\, d\sigma_x(\gamma)\right) d\bar\mu. \tag{3.2}$$

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

By restricting η to suitable subsets of $\mathbb{R}^d \times \Gamma_T$, several manipulations with superposition solutions of the continuity equation are possible and useful, and these are not immediate to see at the level of general solutions of the continuity equation. This is why the following result is interesting (although it plays a little role in these notes, unlike the concept of superposition solution).

Theorem 3.4 (Superposition principle). *Let $\mu_t \in \mathscr{M}_+(\mathbb{R}^d)$ solve PDE and assume that $\int_0^T \|b_t\|_{L^1(\mu_t)}\,dt < +\infty$. Then μ_t is a superposition solution, i.e. there exists $\eta \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t = \mu_t^\eta$ for any $t \in [0,T]$.*

Proof. Here we just a hint of the proof, referring to Chapter 9 of [9] for a detailed one. We mollify μ_t w.r.t. the space variable with a Gaussian kernel ρ (or any other kernel whose support is the whole space), obtaining smooth and strictly positive functions μ_t^ϵ. Defining

$$b_t^\epsilon := \frac{(b_t\mu_t) * \rho_\epsilon}{\mu_t^\epsilon}$$

it is immediate that

$$\dot{\mu}_t^\epsilon + D \cdot (b_t^\epsilon \mu_t^\epsilon) = 0$$

and b_t^ϵ is locally Lipschitz w.r.t. x, with a sufficient uniformity w.r.t. t, so that the representation $\mu_t^\epsilon = X^\epsilon(t,\cdot)_\# \mu_0^\epsilon$. Then, we define

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

so that

$$\int_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t))\,d\eta^\epsilon = \int_{\mathbb{R}^d} \varphi(X^\epsilon(t,x))\,d\mu_0^\epsilon(x) = \int_{\mathbb{R}^d} \varphi\,d\mu_t^\epsilon. \tag{3.3}$$

Since

$$\int_{\mathbb{R}^d \times \Gamma_T} \int_0^T |\dot\gamma|\,dt\,d\eta^\epsilon(x,\gamma) = \int_0^T \int_{\mathbb{R}^d} |b_t^\epsilon|\,d\mu_t^\epsilon\,dt \le \int_0^T \|b_t\|_{L^1(\mu_t)}\,dt$$

and the length functional has compact sublevels in Γ_T, Prokhorov theorem tells us that the family η_ϵ is tight as $\epsilon \downarrow 0$, and if η is any limit point we can pass to the limit in (3.3) to obtain that $\mu_t = \mu_t^\eta$. The more delicate part of the proof is the final one, where it is shown that η is concentrated on the solutions of the ODE. □

The applicability of Theorem 3.1 is strongly limited by the fact that pointwise uniqueness properties are known only in very special situations, for instance when there is a Lipschitz or a one-sided Lipschitz condition. It turns out that in many cases uniqueness of the PDE can only be proved in smaller classes $\mathscr{L}$ of solutions, and it is natural to think that this should reflect into a weaker uniqueness condition at the level of the ODE.

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

$$0 \le \mu'_t \le \mu_t \in \mathscr{L} \qquad \Longrightarrow \qquad \mu'_t \in \mathscr{L}. \tag{3.4}$$

The typical application will be with absolutely continuous measures μ_t, whose densities satisfy some quantitative and possibly time-depending bound.

Definition 3.5 ($\mathscr{L}$-lagrangian flows). *Given the class $\mathscr{L}$, we say that $X(t,x)$ is a $\mathscr{L}$-Lagrangian flow starting from $\bar\mu$ if the following two properties hold:*

(a) $X(\cdot,x)$ solves the ODE for $\bar\mu$-a.e. x;

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

Heuristically $\mathscr{L}$-Lagrangian flows can be thought as suitable selections of the solutions of the ODE (possibly non unique), made in such a way to produce a density in $\mathscr{L}$. The following theorem shows that the $\mathscr{L}$-Lagrangian flow starting from $\bar\mu$ is unique, modulo $\bar\mu$-negligible sets, whenever a comparison principle for the PDE holds, in the class $\mathscr{L}$. We will show existence (and stability) in the Sobolev or BV context in the next sections: in these cases the measure $\bar\mu$ is absolutely continuous w.r.t. $\mathscr{L}^d$.

Theorem 3.6 (Uniqueness). *Assume that the PDE fulfils the comparison principle in $\mathscr{L}$. Then the $\mathscr{L}$-Lagrangian flow starting from $\bar\mu$ is unique, i.e. two different selections $X(t,x)$ and $\tilde X(t,x)$ of solutions of the ODE producing solutions of the PDE in $\mathscr{L}$ satisfy*

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

Proof. If the statement is false we can produce a measure η not concentrated on a graph inducing a solution $\mu^\eta_t \in \mathscr{L}$ of the PDE. This is not possible, thanks to the next result. The measure η can be built as follows:

$$\eta := \frac{1}{2}\Big((x, X(\cdot,x))_\#\bar\mu + (x,\tilde X(\cdot,x))_\#\bar\mu\Big).$$

Since $\mathscr{L}$ is convex we still have $\mu^\eta_t \in \mathscr{L}$. □

Theorem 3.7. *Assume that the PDE fulfils the comparison principle in $\mathscr{L}$. Let $\eta \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ be as in Definition 3.2 and assume that $\mu^\eta_t \in \mathscr{L}$. Then η is concentrated on a graph, i.e. there exists a function $x \mapsto X(\cdot,x) \in \Gamma_T$ such that*

$$\eta = \big(x, \Gamma(\cdot,x)\big)_\#\bar\mu, \quad \text{with} \quad \bar\mu := (\pi_{\mathbb{R}^d})_\#\eta = \mu^\eta_0.$$

Proof. We use the representation (3.2) of η, given by the disintegration theorem, the criterion stated in Lemma 3.9 below and argue by contradiction. If the thesis is false then η_x is not a Dirac mass in a set of $\bar\mu$ positive measure and we can find $t \in (0,T]$, disjoint Borel sets E, $E' \subset \mathbb{R}^d$ and a Borel set C with $\bar\mu(C) > 0$ such that

$$\eta_x(\{\gamma : \ \gamma(t) \in E\})\,\eta_x(\{\gamma : \ \gamma(t) \in E'\}) > 0 \qquad \forall x \in C.$$

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

$$0 < \eta_x(\{\gamma : \ \gamma(t) \in E\}) \leq M\eta_x(\{\gamma : \ \gamma(t) \in E'\}) \qquad \forall x \in C \tag{3.5}$$

for some constant M. We define measure-valued maps η^1, η^2 by

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

and denote by μ^i_t the superposition solutions induced by η^i. Then

$$\mu^1_0 = \eta_{tx}(E)\bar\mu \llcorner C, \qquad \mu^2_0 = M\eta_{tx}(E')\bar\mu \llcorner C,$$

so that (3.5) yields $\mu^1_0 \leq \mu^2_0$. On the other hand

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

Notice also that $\mu^i_t \leq \mu_t$ and so the monotonicity assumption (3.4) on $\mathscr{L}$ gives $\mu^i_t \in \mathscr{L}$. This contradicts the assumption on the validity of the comparison principle in $\mathscr{L}$. □

Exercise 3.8. Let $\sigma \in \mathscr{M}_+(X)$ and let $D \subset [0,T]$ be a dense set. Show that σ is a Dirac mass in Γ_T iff its projections $\gamma(t)_\#\sigma$, $t \in D$, are Dirac masses in $\mathbb{R}^d$.

Lemma 3.9. *Let η_x be a measurable family of positive finite measures in Γ_T with the following property: for any $t \in [0,T]$ and any pair of disjoint Borel sets E, $E' \subset \mathbb{R}^d$ we have*

$$\eta_x(\{\gamma : \ \gamma(t) \in E\})\,\eta_x(\{\gamma : \ \gamma(t) \in E'\}) = 0 \quad \bar\mu\text{-a.e. in } \mathbb{R}^d. \tag{3.6}$$

Then η_x is a Dirac mass for $\bar\mu$-a.e. x.

Proof. Taking into account Exercise 3.8, for a given $t \in (0,T]$ it suffices to check that the measures $\lambda_x := \gamma(t)_\#\eta_x$ are Dirac masses for $\bar\mu$-a.e. x. Then (3.6) gives $\lambda_x(E)\lambda_x(E') = 0$ $\bar\mu$-a.e. for any pair of disjoint Borel sets E, $E' \subset \mathbb{R}^d$. Let $\delta > 0$ and let us consider a partition of $\mathbb{R}^d$ in countably many sets R_i having a diameter less then δ. Then, as $\lambda_x(R_i)\lambda_x(R_j) = 0$ μ-a.e. whenever $i \neq j$, we have a corresponding decomposition of $\bar\mu$-almost all of $\mathbb{R}^d$ in Borel sets A_i such that $\operatorname{supp}\lambda_x \subset \overline{R}_i$ for any $x \in A_i$ (just take $\{\sigma_x(R_i) > 0\}$ and subtract from him all other sets $\{\sigma_x(R_j) > 0\}$, $j \neq i$). Since δ is arbitrary the statement is proved. □

4 The DiPerna–Lions theory

The key ingredient of the theory is the concept of renormalized solution. Before introducing this concept, we define the distribution $b \cdot \nabla w$ in $(0,T) \times \mathbb{R}^d$ as follows

$$\langle b\nabla w, \varphi\rangle := -\int w\langle b, \nabla\varphi\rangle dtdx - \int_0^T \langle D_x \cdot b_t, w_t\varphi_t\rangle\, dt.$$

Notice that this is consistent with the case when w is smooth, and that the second integral makes sense only when $D_x \cdot b_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$.

Definition 4.1 (Renormalized solutions). *Let $b \in L^1_{\rm loc}\left((0,T); L^1_{\rm loc}(\mathbb{R}^d)\right)$ be such that $D \cdot b_t = {\rm div}\, b_t \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$, with*

$$\text{div}\, b_t \in L^1_{\rm loc}\left((0,T); L^1_{\rm loc}(\mathbb{R}^d)\right). \tag{4.1}$$

Let $w \in L^\infty_{\rm loc}\left((0,T); L^\infty(\mathbb{R}^d)\right)$ and assume that

$$c := \frac{d}{dt}w + b \cdot \nabla w \in L^1_{\rm loc}\left((0,T) \times \mathbb{R}^d\right). \tag{4.2}$$

Then, we say that w is a renormalized solution if

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

One of the main results of [24] states that under a Sobolev regularity assumption on b_t any distributional solution is in fact also a renormalized one.

Theorem 4.2 (Renormalization theorem). *Let b as in Definition 4.1 and assume in addition that*

$$b \in L^1_{\rm loc}\left((0,T); W^{1,1}_{\rm loc}(\mathbb{R}^d)\right).$$

Then any distributional solution of (4.2) *is a renormalized solution.*

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

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

to obtain

$$\dot{w}^\epsilon + b \cdot \nabla w^\epsilon = c * \rho_\epsilon + r^\epsilon.$$

By the smoothness of w^ϵ w.r.t. x, the PDE above tells that $\dot{w}^\epsilon_t \in L^1_{\rm loc}$, therefore $w^\epsilon \in W^{1,1}_{\rm loc}\left((0,T) \times \mathbb{R}^d\right)$ and we can apply the standard chain rule in Sobolev spaces, getting

$$\dot{\beta}(w^\epsilon) + b \cdot \nabla\beta(w^\epsilon) = \beta'(w^\epsilon)c * \rho_\epsilon + \beta'(w^\epsilon)r^\epsilon.$$

When we let $\epsilon \downarrow 0$ the convergence in the distribution sense of all terms in the identity above is trivial, with the exception of the last one. To ensure its convergence to zero, it seems necessary to show that $r^\epsilon \to 0$ strongly in $L^1_{\rm loc}$ (remember that $\beta'(w^\epsilon)$ is locally equibounded w.r.t. ϵ). This strong convergence of the "commutators" r^ϵ can be achieved as follows. Playing with the definitions of $b\cdot\nabla w$ and convolution product of a distribution, one proves first the identity

$$r^\epsilon(t,x) = \int_{\mathbb{R}^d} w(t,x-\epsilon y)\frac{b_t(x-\epsilon y)-b_t(x)}{\epsilon}dy - (w{\rm div}\, b_t)*\rho_\epsilon(t,x). \tag{4.3}$$

Then, one uses the strong convergence of translations in L^p and the strong convergence of the difference quotients (a property that characterizes functions in Sobolev spaces)

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

to obtain that r^ϵ strongly converges in $L^1_{\rm loc}$ to

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

The elementary identity

$$\int_{\mathbb{R}^d} y_i\frac{\partial\rho}{\partial y_j}\,dy = -\delta_{ij} \tag{4.4}$$

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

Using the renormalization teorem we can prove a comparison principle in the class $\mathscr{L}$ defined below.

$$\mathscr{L} := \left\{w\in L^\infty\left([0,T];L^1(\mathbb{R}^d)\right)\cap L^\infty\left([0,T];L^\infty(\mathbb{R}^d)\right):\ w\in C\left([0,T];w^*-L^\infty(\mathbb{R}^d)\right)\right\}. \tag{4.5}$$

Theorem 4.3 (Comparison principle). *Assume that*

$$\frac{b}{1+|x|}\in L^1\left([0,T];L^\infty(\mathbb{R}^d)\right)+L^1\left([0,T];\cap L^1(\mathbb{R}^d)\right), \tag{4.6}$$

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

$$\int_0^T \|[{\rm div}\, b_t]^-\|_\infty\,dt < +\infty. \tag{4.7}$$

Assume in addition that any solution of (4.2) *is renormalized. Then the comparison principle holds in the class* $\mathscr{L}$ *defined in* (4.5).

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

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

Notice that β is convex, that $\beta(t) = 0$ iff $t \le 0$ and that $\beta(t) \le t$ for any $t \in \mathbb{R}$. We know that

$$\frac{d}{dt}\beta(w_t) + D_x \cdot \beta(w_t) = \operatorname{div} b_t(\beta(w_t) - w_t\beta'(w_t))$$

in the sense of distributions in $\mathbb{R} \times \mathbb{R}^d$, and clearly $\beta(w_t) = 0$ $\mathscr{L}^{d+1}$-a.e. on $(-\infty, 0) \times \mathbb{R}^d$. Plugging $\varphi_R(\cdot) := \varphi(\cdot/R)$, with $R \ge 1$, into the PDE we obtain

$$\frac{d}{dt}\int_{\mathbb{R}^d} \varphi_R \beta(w_t)\, dx \int_{\mathbb{R}^d} \beta(w_t)\langle b_t, \nabla\varphi_R\rangle\, dx + \int_{\mathbb{R}^d} \varphi_R \operatorname{div} b_t(\beta(w_t) - w_t\beta'(w_t))\, dx.$$

Splitting b as $b_1 + b_2$, with

$$\frac{b_1}{1+|x|} \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right) \quad \text{and} \quad \frac{b_2}{1+|x|} \in L^1\left([0,T]; L^1(\mathbb{R}^d)\right)$$

and using the inequality

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

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

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

The second integral can be estimated with

$$\|[\operatorname{div} b_t]^-\|_\infty \int_{\mathbb{R}^d} \varphi_R \beta(w_t)\, dx,$$

taking into account that $0 \le t\beta'(t) - \beta(t) \le \beta(t)$.

These inequalities have to be understood in the sense of distributions in $\mathbb{R}$, since we don't know a priori if $t \mapsto \beta(w_t)$ is continuous or not. Passing to the limit as $R \to \infty$ and using the integrability assumption on b we get

$$\frac{d}{dt}\int_{\mathbb{R}^d} \beta(w_t)\, dx \le \|[\operatorname{div} b_t]^-\|_\infty \int_{\mathbb{R}^d} \beta(w_t)\, dx$$

in the distribution sense in $\mathbb{R}$. Since the function vanishes for negative times, this suffices to conclude using Gronwall lemma and (4.7). □

5 BV dependence with respect to the spatial variables

One of the main results of [6] is the following one, where we obtain the renormalization lemma under a BV dependence w.r.t. the spatial variables (but still assuming that $D \cdot b_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$).

Theorem 5.1. *Let b as in Definition 4.1 and assume in addition that*

$$b \in L^1_{\mathrm{loc}}\left((0,T); BV_{\mathrm{loc}}(\mathbb{R}^d)\right).$$

Then any distributional solution of (4.2) *is a renormalized solution.*

This section is devoted to a reasonably detailed proof of this result. Before doing that we set up some notation, denoting by

$$Db_t = \nabla b_t \mathscr{L}^d + D^s b_t$$

the Radon–Nikodym decomposition of Db_t in absolutely continuous and singular part w.r.t. $\mathscr{L}^d$. We also introduce the measures $|Db|$ and $|D^s b|$ by integration w.r.t. the time variable, i.e.

$$\int \varphi(t,x)\, d|Db| := \int_0^T \int_{\mathbb{R}^d} \varphi(t,x)\, d|Db_t|\, dt, \qquad \int \varphi(t,x)\, d|D^s b| := \int_0^T \int_{\mathbb{R}^d} \varphi(t,x)\, d|D^s b_t|\, dt.$$

Let us start from the expression (4.3) of the commutators: since $b(t,\cdot) \notin W^{1,1}$ we cannot use the strong convergence of the difference quotients, as done in [24]. However, for any function $u \in BV_{\mathrm{loc}}$ and any $z \in \mathbb{R}^d$ we have the classical L^1 estimate on the difference quotients (see for instance [5])

$$\int_K |u(x+z) - u(x)|\, dx \leq |D_z u|(K_\epsilon) \quad \text{for any } K \subset \mathbb{R}^d \text{ compact},$$

where $Du = (D_1 u, \ldots, D_d u)$ stands for the distributional derivative of u, $D_z u = \langle Du, z\rangle = \sum_i z_i D_i u$ denotes the component along z of Du and K_ϵ is the open ϵ-neighbourhood of K. We notice that

$$D_z \langle b_t, \nabla\rho(z)\rangle = \langle M_t(\cdot) z, \nabla\rho(z)\rangle |Db|$$

and therefore the L^1 estimate on difference quotients gives

$$\limsup_{\epsilon\downarrow 0} \int_K |r^\epsilon|\, dx \leq \|w\|_\infty \int_K \int_{\mathbb{R}^d} |\langle M_t(x) z, \nabla\rho(z)\rangle|\, dz d|Db|(t,x) \tag{5.1}$$

for any compact set $K \subset (0,T) \times \mathbb{R}^d$.

On the other hand, a different estimate of the commutators that reduces to the standard one when $b(t,\cdot)\in W^{1,1}_{\rm loc}$ can be achieved as follows. Let us start from the case $d=1$: if μ is a $\mathbb{R}^m$-valued measure in $\mathbb{R}$ with locally finite variation, then by Fubini's theorem the functions

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

satisfy

$$\int_K|\hat\mu_\epsilon|\,dt\le|\mu|(K_\epsilon)\quad\text{for any compact set } K\subset\mathbb{R}, \tag{5.2}$$

where K_ϵ is the open ϵ neighbourhood of K. A density argument based on (5.2) then shows that $\hat\mu_\epsilon$ converge in $L^1_{\rm loc}(\mathbb{R})$ to the density of μ with respect to $\mathscr{L}^1$ whenever $\mu\ll\mathscr{L}^1$. If $u\in BV_{\rm loc}$ and $\epsilon>0$ we know that

$$\frac{u(x+\epsilon)-u(x)}{\epsilon}=\frac{Du([x,x+\epsilon])}{\epsilon}=\frac{D^au([x,x+\epsilon])}{\epsilon}+\frac{D^su([x,x+\epsilon])}{\epsilon}$$

for $\mathscr{L}^1$-a.e. x (the exceptional set possibly depends on ϵ). In this way we have split the difference quotient, as the sum of two functions, one strongly converging to ∇u in $L^1_{\rm loc}$, and the other one having an L^1 norm on any compact set K asymptotically smaller than $|D^su|(K)$.

If we fix the direction z of the difference quotient, the slicing theory of BV functions (see [5]) gives that this decomposition can be carried on also in d-dimensions, showing that the difference quotients

$$\frac{b_t(x-\epsilon z)-b_t(x)}{\epsilon}$$

can be canonically split into two parts, the first one strongly converging in $L^1_{\rm loc}(\mathbb{R}^d)$ to $\nabla b_t(x)z$, and the second one satisfying having an L^1 norm on K asymptotically smaller than $|\langle D^sb_t,z\rangle|(K)$. Then, repeating the DiPerna–Lions argument and taking into account the error induced by the presence of the second part of the difference quotients we get

$$\limsup_{\epsilon\downarrow0}\int_K|r^\epsilon|\,dx\le\|w\|_\infty\int_K\int_{\mathbb{R}^d}|z||\nabla\rho(z)|\,dzd|D^sb|(t,x) \tag{5.3}$$

for any compact set $K\subset(0,T)\times\mathbb{R}^d$. Roughly speaking, the estimate (5.3) is useful in the regions where the absolutely continuous part is the dominant one, so that $|D^sb|(K)<<1)$, while (5.1) turns out to be useful in the regions where the dominant part is the singular one. Let us see how the two estimates can be combined: coming back to the smoothing scheme, we have

$$\dot\beta(w^\epsilon)+b\cdot\nabla\beta(w^\epsilon)-\beta'(w^\epsilon)c*\rho_\epsilon=\beta'(w_\epsilon)r^\epsilon \tag{5.4}$$

Let us work on an open set $A\subset\subset(0,T)\times\mathbb{R}^d$, let $L=\|w\|_{L^\infty(A)}$ and let L' be the supremum of $|\beta'|$ on $[-M,M]$. Then, (5.3) tells us that limit measure ν of $|\beta'(w^\epsilon)r^\epsilon|\mathscr{L}^d$

as $\epsilon \downarrow 0$ satisfies

$$\nu \llcorner A \le LL' I(\rho)|D^s b| \quad \text{with} \quad I(\rho) := \int_{\mathbb{R}^d} |z||\nabla \rho(z)|\, dz.$$

In particular $\nu \llcorner A$ is singular with respect to $\mathscr{L}^d$. On the other hand, the estimate (5.1) tells also us that

$$\nu \llcorner A \le LL' \int_{\mathbb{R}^d} |\langle M_t(\cdot) z, \nabla \rho(z) \rangle|\, dz |Db|.$$

These two estimates imply that

$$\nu \llcorner A \le LL' \int_{\mathbb{R}^d} |\langle M_t(\cdot) z, \nabla \rho(z) \rangle|\, dz |D^s b|. \tag{5.5}$$

Notice that in this way we got rid of the potentially dangerous term $I(\rho)$: in fact, we are going to choose very anisotropic kernels ρ on which $I(\rho)$ can be arbitrarily large. The measure ν can of course depend on the choice of ρ, but (5.4) tells us that the measure

$$\sigma := \frac{d}{dt}\beta(w_t) + b \cdot \nabla \beta(w_t) - e_t w_t \beta'(w_t),$$

clearly independent of ρ, satisfies $|\sigma| \le \nu$ in A. Eventually we obtain

$$|\sigma| \llcorner A \le LL' \Lambda(M_\cdot(\cdot), \rho)|D^s b| \quad \text{with} \quad \Lambda(N, \rho) := \int_{\mathbb{R}^d} |\langle Nz, \nabla \rho(z) \rangle|\, dz. \tag{5.6}$$

We are thus led to the minimum problem

$$G(N) := \inf \left\{ \Lambda(N, \rho) : \ \rho \in C_c^\infty(B_1),\ \rho \ge 0,\ \int_{\mathbb{R}^d} \rho = 1 \right\} \tag{5.7}$$

with $N = M_t(x)$. Notice that (5.6) gives

$$|\sigma| \llcorner A \le LL' \inf_{\rho \in D} \Lambda(M_\cdot(\cdot), \rho)|D^s b|$$

for any countable set D of kernels ρ, and the continuity of $\rho \mapsto \Lambda(N, \rho)$ w.r.t. the $W^{1,1}(B_1)$ norm and the separability of $W^{1,1}(B_1)$ give

$$|\sigma| \llcorner A \le LL' G(M_\cdot(\cdot))|D^s b|. \tag{5.8}$$

Notice now that the assumption that $D \cdot b_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in (0, T)$ gives

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

Hence, recalling the definition of $|D^s b|$, the trace of $M_t(x)$ vanishes for $|D^s b|$-a.e. (t, x). Applying the following lemma, due to Alberti [?] [1] and using (5.8) we conclude that $\sigma = 0$, concluding the proof.

[1] Actually this lemma came out during the Luminy school in October of 2003, where I raised the problem of computing the infimum in (5.7), and Alberti came up with the solution!

Lemma 5.2 (Alberti). *For any $d \times d$ matrix N the infimum in* (5.7) *is* $|\text{trace}\, N|$.

Proof. Notice first that the lower bound follows immediately by the identity

$$\int_{\mathbb{R}^d} \langle Nz, \nabla \rho(z) \rangle \, dz = -\text{trace}\, N,$$

that in turn follows by (4.4). Hence, we have to show only the upper bound. Since

$$\langle Nz, \nabla \rho(z) \rangle = \text{div}\, (Nz\rho(z)) - \text{trace}\, N\rho(z)$$

it suffices to show that for any $T > 0$ there exists ρ such that

$$\int_{\mathbb{R}^d} |\text{div}\, (Nz\rho(z))| \, dz \le \frac{2}{T}. \tag{5.9}$$

The heuristic idea is to build ρ as the superposition of elementary probability measures associated to the curves $e^{tN}x$, $0 \le t \le T$, on which the divergence operator can be easily estimated. Given a smooth kernel θ with compact support, it turns out that the function

$$\rho(z) := \frac{1}{T} \int_0^T \theta(e^{-tN}z) e^{-t\, \text{trace}\, N} \, dt \tag{5.10}$$

has the required properties (here $e^{tN}x = \sum_i t^i N^i x / i!$ is the solution of the ODE $\dot{\gamma} = N\gamma$ with the initial condition $\gamma(0) = x$). Indeed, it is immediate to check that ρ is smooth and compactly supported. To estimate the divergence of $Nz\rho(z)$, we notice that $\rho = \int \theta(x)\mu_x \, dx$, where μ_x are the probability 1-dimensional measures concentrated on the image of the curves $t \mapsto e^{tN}x$ defined by

$$\mu_x := (e^{\cdot N}x)_\#(\frac{1}{T}\mathscr{L}^1 \llcorner [0, T]).$$

Indeed, for any $\varphi \in C_c^\infty(\mathbb{R}^d)$ we have

$$\begin{aligned} \int_{\mathbb{R}^d} \theta(x) \langle \mu_x, \varphi \rangle \, dx &= \frac{1}{T} \int_0^T \int_{\mathbb{R}^d} \theta(x) \varphi(e^{tN}x) \, dt dx \\ &= \frac{1}{T} \int_0^T \int_{\mathbb{R}^d} \theta(e^{-tN}y) e^{-t\, \text{traceN}} \varphi(y) \, dy dt = \int_{\mathbb{R}^d} \rho(y)\varphi(y) \, dy. \end{aligned}$$

By the linearity of the divergence operator, it suffices to check that

$$|D \cdot (Nz\mu_x)| \le \frac{2}{T} \qquad \forall x \in \mathbb{R}^d.$$

But this is elementary, since

$$\int_{\mathbb{R}^d} \langle Nz, \nabla \varphi(z) \rangle \, d\mu_x(z) = \frac{1}{T} \int_0^T \langle Ne^{tN}x, \nabla \varphi(e^{tN}x) \rangle \, dt = \frac{\varphi(e^{TN}x) - \varphi(x)}{T}$$

for any $\varphi \in C_c^\infty(\mathbb{R}^d)$, so that $TD \cdot (Nx\mu_x) = \delta_x - \delta_{e^{TN}x}$. □

We conclude this section noticing that the original argument in [6] is slightly different and uses, instead of Lemma 5.2, a much deeper result, still due to Alberti [2], saying that for a $BV_{\rm loc}$ function $u : \mathbb{R}^d \to \mathbb{R}^m$ the matrix $M(x)$ in the polar decomposition $Du = M|Du|$ has rank 1 for $|D^s u|$-a.e. x, i.e. there exist unit vectors $\xi(x) \in \mathbb{R}^n$ and $\eta(x) \in \mathbb{R}^m$ such that $M(x)z = \eta(x)\langle z, \xi(x)\rangle$. In this case the asymptotically optimal kernels are easy to build, by mollifying in the ξ direction much faster than in all other directions (this is precisely what Bouchut did in [12]).

6 Existence and stability of Lagrangian flows

In this section we study the Lagrangian counterpart of the well-posedness results obtained in the previous two sections, in the Sobolev and in the BV case.

Theorem 6.1 (Existence). *Let $b \in L^1_{\rm loc}\left((0,T); BV_{\rm loc}(\mathbb{R}^d)\right)$ be satisfying (4.6) and let us assume that $D_x \cdot b_t = {\rm div}\, b_t \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in [0,T]$, with $\|[{\rm div}\, b_t]^-\|_\infty \in L^1(0,T)$. Let $\mathscr{L}$ be defined as in (4.5) and assume that the comparison principle holds in $\mathscr{L}^+$. Then for any $\bar\mu = \rho\mathscr{L}^d$ with a nonnegative $\rho \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ there exists a $\mathscr{L}^d$-Lagrangian flow starting from $\bar\mu$. Moreover, the flow satisfies*

$$X(t,\cdot)_\#\bar\mu \le \|\rho\|_\infty \exp\left(\int_0^t \|[{\rm div}\, b_s]^-\|_\infty\, ds\right)\mathscr{L}^d. \tag{6.1}$$

Proof. **Step 1.** (Smoothing) We define $b^\epsilon = b * \rho_\epsilon$ by mollifying with respect to the spatial variable with a kernel ρ with compact support, with $\epsilon \in (0,1)$. Notice that

$$\int_0^T \|[{\rm div}\, b^\epsilon_t]^-\|_\infty\, dt \le \int_0^T \|[{\rm div}\, b_t]^-\|_\infty\, dt. \tag{6.2}$$

Splitting $b = (b_1 + b_2)(1+|x|)$ according to (4.6), with

$$b_1 \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right), \qquad b_2 \in L^1\left([0,T]; L^1(\mathbb{R}^d)\right),$$

we notice that

$$((1+|x|)b_i) * \rho_\epsilon \le 2(1+|x|)(|b_i| * \rho_\epsilon) \qquad i = 1,2$$

and therefore $b^\epsilon \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right)$ and we can also write $b^\epsilon = b^\epsilon_1 + b^\epsilon_2$ with

$$\sup_{\epsilon\in(0,1)} \|\frac{b^\epsilon_1}{1+|x|}\|_{L^1((0,T);L^\infty(\mathbb{R}^d))} < +\infty, \qquad \sup_{\epsilon\in(0,1)} \|\frac{b^\epsilon_2}{1+|x|}\|_{L^1((0,T);L^1(\mathbb{R}^d))} < +\infty, \tag{6.3}$$

Therefore we can apply Remark 2.2 and consider the characteristics $X^\epsilon(t,x)$ associated to b^ϵ, together with the induced measures

$$\mu^\epsilon_t := X^\epsilon(t,\cdot)_\#\bar\mu, \qquad \eta_\epsilon := (x, X(\cdot,x))_\#\,\bar\mu$$

with $\mu_t^\epsilon = \mu_t^{\eta_\epsilon}$. Notice also that the explicit representation (2.3) gives

$$\mu_t^\epsilon = w_t^\epsilon \mathscr{L}^d \quad \text{with} \quad w_t^\epsilon = \frac{\rho}{JX^\epsilon(t,\cdot)} \circ (X^\epsilon(t,\cdot))^{-1} \tag{6.4}$$

and (2.4) together with (6.2) give

$$w_t^\epsilon \le \|\rho\|_\infty \exp\left(\int_0^T \|[\operatorname{div} b_s]^-\|_\infty \, ds\right). \tag{6.5}$$

Step 2. (Tightness of η_ϵ) We claim that the family η_ϵ is tight: indeed, by the remarks made after Definition 8.2, it suffices to find a coercive functional $\Psi : \mathbb{R}^d \times \Gamma_T \to [0, +\infty)$ whose integral w.r.t. all measures η_ϵ is uniformly bounded. Since $\bar\mu$ has finite mass we can find a function $\varphi : \mathbb{R}^d \to [0, +\infty)$ such that $\varphi \in L^1(\bar\mu)$ and $\varphi(x) \to +\infty$ as $|x| \to \infty$. Then, we define

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

and notice that the coercivity of Ψ follows immediately from Lemma 6.2 below. Then, we compute:

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

Using the uniform $L^\infty\left((0,T); L^\infty(\mathbb{R}^d)\right)$ estimate on w_t^ϵ given by (6.5), the uniform $L^\infty\left((0,T); L^1(\mathbb{R}^d)\right)$ estimate coming from the fact that w_t^ϵ are probability densities and the uniform estimate (6.3) we conclude that the integrals of Ψ are uniformly bounded.

Step 3. (The limit flow belongs to $\mathscr{L}$) Let now η be a narrow limit point of η_ϵ along some infinitesimal sequence ϵ_i and write $\eta_i = \eta_{\epsilon_i}$ in short. Let us show first that the induced flow μ_t^η belongs to $\mathscr{L}$. Since $\mu_t^{\eta_i}$ narrowly converge to μ_t^η as $i \to \infty$ it is immediate to infer from (6.4) and (6.5) that

$$\mu_t^\eta = w_t \mathscr{L}^d \quad \text{with} \quad w_t \le \|\rho\|_\infty \exp\left(\int_0^t \|[\operatorname{div} b_s]^-\|_\infty \, ds\right). \tag{6.6}$$

Therefore the narrow continuity of $t \mapsto \mu_t^\eta$ immediately yields the w^*-continuity of $t \mapsto w_t$ and this proves that $\mu_t^\eta \in \mathscr{L}$.

Step 4. (η is concentrated on solutions of the ODE) Next we show that η is concentrated on the class of solutions of the ODE. Let $\bar t \in [0,T]$, $\chi \in C_c^\infty(\mathbb{R}^d)$ with $0 \le \chi \le 1$, $c \in L^1\left([0,\bar t]; L^\infty(\mathbb{R}^d)\right)$, with $c(t,\cdot)$ continuous in $\mathbb{R}^d$ for $\mathscr{L}^1$-a.e. $t \in [0,T]$, and define

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

In the following we use repeatedly this fact, whose proof immediately follows by the dominated convergence theorem: if

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

is bounded and converges to $c/(1+|x|)$ $\mathscr{L}^{d+1}$-a.e., then

$$\lim_{h\to\infty} \int_0^T \int_{\mathbb{R}^d} \frac{|c_h - c|}{1+|x|^{d+2}}\,dxdt = 0. \tag{6.7}$$

It is immediate to check that $\Phi_c^{\bar t} \in C_b(\mathbb{R}^d \times \Gamma_T)$, so that using (6.7), the fact that $b^\epsilon/(1+|x|)$ in bounded in $L^1(L^1) + L^1(L^\infty)$ and (6.4), (6.5) we get

$$\begin{aligned}
\int_{\mathbb{R}^d\times\Gamma_T} \Phi_c^{\bar t}\,d\eta &= \lim_{i\to\infty} \int_{\mathbb{R}^d\times\Gamma_T} \Phi_c^{\bar t}\,d\eta_i \\
&= \lim_{i\to\infty} \int_{\mathbb{R}^d} \chi(x)\rho(x)\frac{\left|\int_0^{\bar t} b^{\epsilon_i}(s,X^{\epsilon_i}(s,x)) - c(s,X^{\epsilon_i}(s,x))\,ds\right|}{1+\sup\limits_{[0,\bar t]}|X^{\epsilon_i}(x,\cdot)|^{d+2}}\,dx \\
&\le \limsup_{i\to\infty} \int_{\mathbb{R}^d}\int_0^{\bar t} \rho(x)\frac{|b^{\epsilon_i}(s,X^{\epsilon_i}(s,x)) - c(s,X^{\epsilon_i}(s,x))|}{1+|X^{\epsilon_i}(s,x)|^{d+2}}\,dsdx \\
&\le \|\rho\|_\infty \exp\left(\int_0^T \|[\operatorname{div} b_s]^-\|_\infty\,ds\right) . \limsup_{i\to\infty} \int_0^{\bar t}\int_{\mathbb{R}^d} \frac{|b^{\epsilon_i}(s,x) - c(s,x)|}{1+|x|^{d+2}}\,dsdx \\
&\le \|\rho\|_\infty \exp\left(\int_0^T \|[\operatorname{div} b_s]^-\|_\infty\,ds\right) \cdot \int_0^{\bar t}\int_{\mathbb{R}^d} \frac{|b(s,x) - c(s,x)|}{1+|x|^{d+2}}\,dsdx.
\end{aligned}$$

In the previous estimate we can now choose $c = b^{\epsilon_i}$ and use (6.7) to obtain

$$\liminf_{i\to\infty} \int_{\mathbb{R}^d\times\Gamma_T} \chi(x)\frac{\left|\gamma(\bar t) - x - \int_0^{\bar t} b^{\epsilon_i}(s,\gamma(s))\,ds\right|}{1+\sup\limits_{[0,\bar t]}|\gamma|^{d+2}}\,d\eta = 0. \tag{6.8}$$

Now, using te upper bound (6.6) and (6.7) again we get

$$
\begin{aligned}
&\liminf_{i\to\infty}\int_{\mathbb{R}^d\times\Gamma_T}\frac{\int_0^{\bar t}|b^{\epsilon_i}(s,\gamma(s))-b(s,\gamma(s))|\,ds}{1+\sup\limits_{[0,\bar t]}|\gamma|^{d+2}}\,d\eta \qquad (6.9)\\
\leq\ &\liminf_{i\to\infty}\int_0^T\int_{\mathbb{R}^d\times\Gamma_T}\frac{|b^{\epsilon_i}(s,\gamma(s))-b(s,\gamma(s))|}{1+|\gamma(s)|^{d+2}}\,d\eta\,ds\\
\leq\ &\|\rho\|_\infty\exp\left(\int_0^T\|[\operatorname{div} b_s]^-\|_\infty\,ds\right)\liminf_{i\to\infty}\int_0^T\int_{\mathbb{R}^d}\frac{|b^{\epsilon_i}-b|}{1+|x|^{d+2}}\,dxds=0.
\end{aligned}
$$

Hence, from (6.8) and (6.9) and Fatou's lemma we infer that for $\chi\eta$-a.e. (x,γ) there is a subsequence $\epsilon_{i(l)}$ such that

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

so that

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

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

Step 5. (Conclusion) Recall also that by Theorem 5.1 any distributional solution of PDE in $\mathscr{L}$ is renormalized, and that Theorem 4.3 ensures as a consequence the comparison principle in $\mathscr{L}$. We are now in the position of applying Theorem 3.7, saying that under these conditions necessarily

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

for a suitable map $x\mapsto X(\cdot,x)$. Clearly, by the concentration property of η, $X(\cdot,x)$ has to be a solution of the ODE for $\bar\mu$-a.e. x. This proves that $X(t,x)$ is a $\mathscr{L}$-Lagrangian flow. □

Lemma 6.2 (A coercive functional in Γ_T). *Let $\varphi:\mathbb{R}^d\to\mathbb{R}$ and let*

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

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

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

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

one obtains that also $\sup_{[0,T]}|\gamma_h|$ is uniformly bounded. As a consequence the factor $1/(1+|\gamma_h|)$ inside the integral part of Φ can be uniformly estimated from below, and therefore the sequence is bounded in $W^{1,1}\left((0,T);\mathbb{R}^d\right)$. The compactness of the embedding of $W^{1,1}\left((0,T);\mathbb{R}^d\right)$ in Γ_T gives the conclusion. □

Theorem 6.3 (Stability of $\mathscr{L}$-Lagrangian flows). *Let b_h, b and $\bar\mu$ be as in Theorem 6.1 and let X_h, X be the corresponding $\mathscr{L}$-Lagrangian flows. Assume that*

(i) $b_h/(1+|x|)$ *is uniformly bounded in*

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

(ii) $\sup_h\int_0^T\|[\operatorname{div} b_{ht}]^-\|_\infty\,dt<+\infty$.

Then

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

Proof. We define η_h as the push forward of $\bar\mu$ under the map $x\mapsto(x,X_h(\cdot,x))$ and argue exactly as in the proof of Theorem 6.1 (using the a-priori upper bound (6.1) and assumptions (i), (ii)) to show that η_h is tight, to find a subsequence (not relabelled) narrowly converging to η in $\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$, to show using (iii) that η is concentrated on the solutions of ODE relative to the limit vector field b. In fact, in all steps of the previous theorem we never used the smoothness of the flows, but only the apriori bounds (6.5) that in this case can be replaced by (6.1), so that the proof applies to families of $\mathscr{L}$-Lagrangian flows as well.

Then Theorem 3.7 says that η is the push-forward under the map $x\mapsto(x,X(\cdot,x))$ of $\bar\mu$, where X is the $\mathscr{L}$-Lagrangian flow starting from $\bar\mu$. Therefore

$$(x,X_h(\cdot,x))_\#\bar\mu \text{ converge narrowly to } (x,X(\cdot,x))_\#\bar\mu.$$

Applying Lemma 8.3 we conclude. □

7 An application to the Keyfitz–Kranzer system of conservation laws

8 Appendix: some basic facts of Measure Theory

In these notes all spaces under consideration are metric and separable, and the σ-algebra involved are the Borel σ-algebras.

Definition 8.1 (Push-forward). *Let $\mu \in \mathscr{M}(X)$ and let $f : X \to Y$ be a Borel map. The push-forward $f_\# \mu \in \mathscr{M}(Y)$ is defined by $f_\# \mu(B) = \mu(f^{-1}(B))$ for any Borel set $B \subset Y$.*

A simple approximation by simple functions shows the chain-rule

$$\int_Y g \, df_\# \mu = \int_X g \circ f \, d\mu \tag{8.1}$$

for any bounded Borel function (or even either nonnegative or nonpositive, and $\overline{\mathbb{R}}$-valued) function $g : Y \to \mathbb{R}$.

Definition 8.2 (Narrow convergence and compactness). *Narrow convergence in $\mathscr{M}(X)$ is the convergence with respect to the duality with $C_b(X)$, the space of continuous and bounded functions in X. By Prokhorov theorem, a bounded set $\mathscr{F}$ in $\mathscr{M}(X)$ is sequentially relatively compact with respect to the narrow convergence if and only if it tight, i.e. for any $\epsilon > 0$ there exists a compact set $K \subset X$ such that $|\mu|(X \setminus K) < \epsilon$ for any $\mu \in \mathscr{F}$.*

It is immediate to check that a sufficient condition for tightness of a bounded family $\mathscr{F}$ is the existence of a *coercive* functional $\Phi : X \to [0, +\infty]$ (i.e. a functional such that its sublevel sets $\{\Phi \le t\}$, $t \in \mathbb{R}^+$, are relatively compact) such that

$$\int_X \Phi(x) d\mu \le 1 \qquad \forall \mu \in \mathscr{F}.$$

Recall that a Y-valued sequence (v_h) is said to converge in μ-measure to v if

$$\lim_{h \to \infty} \mu\left(\{d_Y(v_h, v) > \delta\}\right) = 0 \qquad \forall \delta > 0.$$

This is equivalent to the L^1 convergence to 0 of the maps $1 \wedge d_Y(v_h, v)$.

Lemma 8.3. *Let v_h, $v : X \to Y$ be Borel maps and let $\mu \in \mathscr{P}(X)$. Then $v_h \to v$ in μ-measure iff*

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

Proof. If $v_h \to v$ in μ-measure then $\varphi(x, v_h(x))$ converges in $L^1(\mu)$ to $\varphi(x, v(x))$, and therefore thanks to (8.1) we immediately obtain the convergence of the push-forward. Conversely, let $\delta > 0$ and, for any $\epsilon > 0$, let $w \in C_b(X; Y)$ such that $\mu(\{v \neq w\}) \leq \epsilon$. We define

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

and notice that

$$\int_{X \times Y} \varphi \, d(x, v_h(x))_\# \mu \geq \mu(\{d_Y(w, v_h) > \delta\}), \quad \int_{X \times Y} \varphi \, d(x, v(x))_\# \mu \leq \mu(\{w \neq v\}).$$

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

$$\limsup_{h \to \infty} \mu(\{d_Y(v, v_h) > \delta\}) \leq 2\mu(\{w \neq v\}) \leq 2\epsilon$$

and since ϵ is arbitrary the proof is achieved. □

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

[6] L.Ambrosio: *Transport equation and Cauchy problem for* BV *vector fields.* Preprint 2003, to appear on Inventiones Mathematicae.

[7] 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.

[8] 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).

[9] L.Ambrosio, N.Gigli, G.Savaré: *Gradient flows in metric spaces and in the Wasserstein space of probability measures.* Book in preparation, to be published by Birkhäuser.

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

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

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

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

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

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

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

[22] C.Dafermos:

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

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

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

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

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

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

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

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

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

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