Transport equation and Cauchy problem for non-smooth vector fields

Luigi Ambrosio

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

Cetraro, June 27–July 2, 2005

Contents

1 Introduction

In these lectures we study the well-posedness of the Cauchy problem for the homogeneous conservative continuity equation

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

and for the transport equation

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

Here $\boldsymbol{b}(t,x) = \boldsymbol{b}_t(x)$ is a given time-dependent vector field in $\mathbb{R}^d$: we are interested to the case when $\boldsymbol{b}_t(\cdot)$ is not necessarily Lipschitz and has, for instance, a Sobolev or BV regularity. Vector fields with this "low" regularity show up, for instance, in several PDE's describing the motion of fluids, and in the theory of conservation laws.
We are also particularly interested to the well posedness of the system of ordinary differential equations

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

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

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

The following simple example provides an illustration of the kind of phenomena that can occur.

Example 1 Let us consider the autonomous ODE

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

Then, solutions of the ODE are not unique for $x_0 = -c^2 < 0$. Indeed, they reach the origin in time $2c$, where can stay for an arbitrary time T, then continuing as $x(t) = \frac{1}{4}(t - T - 2c)^2$. Let us consider for instance the Lipschitz approximation (that could easily be made smooth) of $b(\gamma) = \sqrt{|\gamma|}$ by

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

with $\lambda_\epsilon - \epsilon^2 > 0$. Then, solutions of the approximating ODE's starting from $-c^2$ reach the value $-\epsilon^2$ in time $t_\epsilon = 2(c-\epsilon)$ and then they continue with constant speed ϵ until they reach $\lambda_\epsilon - \epsilon^2$, in time $T_\epsilon = \lambda_\epsilon/\epsilon$. Then, they continue as $\lambda_\epsilon - 2\epsilon^2 + \frac{1}{4}(t - t_\epsilon - T_\epsilon)^2$.
Choosing $\lambda_\epsilon = \epsilon T$, with $T > 0$, by this approximation we select the solutions that don't move, when at the origin, exactly for a time T.

Other approximations, as for instance $b_\epsilon(\gamma) = \sqrt{\epsilon + |\gamma|}$, select the solutions that move immediately away from the singularity at $\gamma = 0$. Among all possibilities, this family of solutions $x(t, x_0)$ is singled out by the property that $x(t,\cdot)_\#\mathscr{L}^1$ is absolutely continuous with respect to $\mathscr{L}^1$, so no concentration of trajectories occurs at the origin. To see this fact, notice that we can integrate in time the identity

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

and use Fubini's theorem to obtain

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

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

We will see that there is a close link between (PDE) and (ODE), first investigated in a nonsmooth setting by Di Perna and Lions in [53].
Let us now make some basic technical remarks on the continuity equation and the transport equation:

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

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

However, when we consider possibly singular measures μ_t, we must take care of the fact that the product $\boldsymbol{b}_t\mu_t$ is sensitive to modifications of $\boldsymbol{b}_t$ in $\mathscr{L}^d$-negligible sets. In the Sobolev or BV case we will consider only measures $\mu_t = w_t\mathscr{L}^d$, so everything is well posed.
(2) On the other hand, due to the fact that the distribution $\boldsymbol{b}_t \cdot \nabla w$ is defined by

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

(a definition consistent with the case when w_t is smooth) the transport equation makes sense *only* if we assume that $D_x \cdot \boldsymbol{b}_t = \operatorname{div} \boldsymbol{b}_t\mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in I$. See also [28], [31] for recent results on the transport equation when $\boldsymbol{b}$ satisfies a one-sided Lipschitz condition.

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

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

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

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

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

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

This duality method is a classical tool to prove uniqueness in a sufficiently smooth setting (but see also [28], [31]).
(3) Finally, if we denote by $\boldsymbol{Y}(t,s,x)$ the solution of the ODE at time t, starting from x at the initial times s, *i.e.*

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

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

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

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

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

This property is used in a essential way in [53] to characterize the flow $\boldsymbol{Y}$ and to prove its stability properties. The approach developed here, based on [7], is based on a careful analysis of the measures transported by the flow, and ultimately on the homogeneous continuity equation only.
Acknowledgement. I wish to thank Gianluca Crippa and Alessio Figalli for their careful reading of a preliminary version of this manuscript.

2 Transport equation and continuity equation within the Cauchy-Lipschitz framework

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

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

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

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

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

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

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

The representation formula is indeed very simple:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

for any Borel set $A \subset \mathbb{R}^d$, 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 that $|Jf| > 0$ $\mathscr{L}^d$-a.e. on A we can set $A = f^{-1}(B)$ and $g = \rho/|Jf|$ to obtain

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

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

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

In our case $f(x) = \boldsymbol{X}(t,x)$ is surely 1-1, onto and Lipschitz. It remains to show that $|J\boldsymbol{X}(t,\cdot)|$ does not vanish: in fact, one can show that $J\boldsymbol{X} > 0$ and

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3 ODE uniqueness versus PDE uniqueness

In this section we illustrate some quite general principles, whose application may depend on specific assumptions on $\boldsymbol{b}$, relating the uniqueness of the ODE to the uniqueness of the PDE. The viewpoint adopted in this section is very close in spirit to Young's theory [85] of generalized surfaces and controls (a theory with remarkable applications also non-linear PDE's [52, 78] and Calculus of Variations [19]) and has also some connection with Brenier's weak solutions of incompressible Euler equations [24], with Kantorovich's viewpoint in the theory of optimal transportation [57, 76] and with Mather's theory [71, 72, 18]: in order to study existence, uniqueness and stability with respect to perturbations of the data of solutions to the ODE, we consider suitable measures in the space of continuous maps, allowing for superposition of trajectories. Then, in some special situations we are able to show that this superposition actually does not occur, but still this "probabilistic" interpretation is very useful to understand the underlying techniques and to give an intrinsic characterization of the flow.
The first very general criterion is the following.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The absolute continuity of the integral shows that the right hand side can be made small when $\sum_i (b_i - a_i)$ is small. This proves the absolute continuity.
It remains to evaluate the time derivative of $t \mapsto \langle \mu_t^{\boldsymbol{\eta}}, \varphi\rangle$: we know that for $\boldsymbol{\eta}$-a.e. (x, γ) the identity $\dot{\gamma}(t) = \boldsymbol{b}_t(\gamma(t))$ is fulfilled for $\mathscr{L}^1$-a.e. $t \in [0, T]$. Then, by Fubini's theorem, we know also that for $\mathscr{L}^1$-a.e. $t \in [0, T]$ the previous identity holds for $\boldsymbol{\eta}$-a.e. (x, γ), and therefore

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

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

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

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

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

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

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

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

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

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

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

Proof. **Step 1 (smoothing).** [58] We mollify μ_t w.r.t. the space variable with a kernel ρ having finite first moment M and support equal to the whole of $\mathbb{R}^d$ (a Gaussian, for instance), obtaining smooth and strictly positive functions μ_t^ϵ. We also choose a function $\psi : \mathbb{R}^d \to [0,+\infty)$ such that $\psi(x) \to +\infty$ as $|x| \to +\infty$ and

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

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

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

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

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

it is immediate that

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

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

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

so that

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

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

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

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

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

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

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

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

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

and

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

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

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

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

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

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

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

and noticing that

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

we can pass to the limit in (13) with $\boldsymbol{c} = \boldsymbol{c}^n$ to obtain (12).
It remains to show (13). This is a limiting argument based on the fact that (12) holds for $\boldsymbol{b}^\epsilon$, $\boldsymbol{\eta}^\epsilon$:

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

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

The applicability of Theorem 9 is strongly limited by the fact that, on one hand, *pointwise* uniqueness properties for the ODE are known only in very special situations, for instance when there is a Lipschitz or a one-sided Lipschitz (or log-Lipschitz, Osgood...) condition on $\boldsymbol{b}$. On the other hand, also uniqueness for general measure-valued solutions is known only in special

situations. It turns out that in many cases uniqueness of the PDE can only be proved in smaller classes $\mathscr{L}$ of solutions, and it is natural to think that this should reflect into a weaker uniqueness condition at the level of the ODE.
We will see indeed that there is uniqueness in the "selection sense". In order to illustrate this concept, in the following we consider a convex class $\mathscr{L}_{\boldsymbol{b}}$ of measure-valued solutions $\mu_t \in \mathscr{M}_+(\mathbb{R}^d)$ of the continuity equation relative to $\boldsymbol{b}$, satifying the following monotonicity property:

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

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

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

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

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

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

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

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

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

Heuristically $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows can be thought as suitable selections of the solutions of the ODE (possibly non unique), made in such a way to produce a density in $\mathscr{L}_{\boldsymbol{b}}$, see Example 1 for an illustration of this concept.
We will show that the $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar\mu$ is unique, modulo $\bar\mu$-negligible sets, whenever a comparison principle for the PDE holds, in the class $\mathscr{L}_{\boldsymbol{b}}$ (*i.e.* the inequality between two solutions at $t=0$ is preserved at later times).
Before stating and proving the uniqueness theorem for $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows, we state two elementary but useful results. The first one is a simple exercise:

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

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

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

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

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

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

Theorem 16 (Uniqueness of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows) *Assume that the PDE fulfils the comparison principle in $\mathscr{L}_{\boldsymbol{b}}$. Then the $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar\mu$ is unique, i.e. two different selections $\boldsymbol{X}_1(t,x)$ and $\boldsymbol{X}_2(t,x)$ of solutions of the ODE inducing solutions of the the continuity equation in $\mathscr{L}_{\boldsymbol{b}}$ satisfy*

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

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

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

Since $\mathscr{L}_{\boldsymbol{b}}$ is convex we still have $\mu_t^{\boldsymbol{\eta}} = \frac{1}{2}(\mu_t^{\boldsymbol{\eta}^1} + \mu_t^{\boldsymbol{\eta}^2}) \in \mathscr{L}_{\boldsymbol{b}}$. □

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

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

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

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

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

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

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

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

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

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

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

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

and denote by μ^i_t the (superposition) solutions of the continuity equation induced by $\boldsymbol{\eta}^i$. Then

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

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

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

Notice also that $\mu^i_t\le\mu_t$ and so the monotonicity assumption (14) on $\mathscr{L}_{\boldsymbol{b}}$ gives $\mu^i_t\in\mathscr{L}_{\boldsymbol{b}}$. This contradicts the assumption on the validity of the comparison principle in $\mathscr{L}_{\boldsymbol{b}}$. □

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

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

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

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

Pushing both sides via e_t we obtain

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

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

Finally, let us discuss the *stability* issue. This is particularly relevant, as we will see, in connection with the applications to PDE's.

Definition 20 (Convergence of velocity fields) We define the convergence of $\boldsymbol{b}^h$ to $\boldsymbol{b}$ in a indirect way, defining rather a convergence of $\mathscr{L}_{\boldsymbol{b}^h}$ to $\mathscr{L}_{\boldsymbol{b}}$: we require that

$$\boldsymbol{b}^h\mu^h_t\rightharpoonup\boldsymbol{b}\mu_t\ \text{ in }(0,T)\times\mathbb{R}^d\qquad\text{and}\qquad\mu_t\in\mathscr{L}_{\boldsymbol{b}}$$

whenever $\mu^h_t\in\mathscr{L}_{\boldsymbol{b}^h}$ and $\mu^h_t\to\mu_t$ narrowly for all $t\in[0,T]$.

For instance, in the typical case when $\mathscr{L}$ is bounded and closed, w.r.t the weak* topology, in $L^\infty(L^1)\cap L^\infty(L^\infty)$, and

$$\mathscr{L}_{\boldsymbol{c}} := \mathscr{L} \cap \left\{ w :\ \frac{d}{dt}w + D_x\cdot(\boldsymbol{c}w) = 0 \right\}$$

the implication is fulfilled whenever $\boldsymbol{b}^h \to \boldsymbol{b}$ strongly in $L^1_{\rm loc}$.
The natural convergence for the stability theorem is *convergence in measure*. Let us recall that a Y-valued sequence (v_h) is said to converge in $\bar\mu$-measure to v if

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

This is equivalent to the L^1 convergence to 0 of the $\mathbb{R}^+$-valued maps $1\wedge d_Y(v_h,v)$.
Recall also that convergence $\bar\mu$-a.e. implies convergence in measure, and that the converse implication is true passing to a suitable subsequence.

Theorem 21 (Stability of $\mathscr{L}$-Lagrangian flows) *Assume that*

(i) $\mathscr{L}_{\boldsymbol{b}^h}$ *converge to* $\mathscr{L}_{\boldsymbol{b}}$*;*

(ii) $\boldsymbol{X}^h$ *are* $\mathscr{L}_{\boldsymbol{b}^h}$*-flows relative to* $\boldsymbol{b}^h$ *starting from* $\bar\mu\in\mathscr{M}_+(\mathbb{R}^d)$ *and* $\boldsymbol{X}$ *is the* $\mathscr{L}_{\boldsymbol{b}}$*-flow relative to* $\boldsymbol{b}$ *starting from* $\bar\mu$*;*

(iii) setting $\mu_t^h := \boldsymbol{X}^h(t,\cdot)_\#\bar\mu$*, we have*

$$\mu_t^h \to \mu_t \qquad \textit{narrowly as } h\to\infty \textit{ for all } t\in[0,T] \tag{17}$$

$$\limsup_{h\to\infty}\int_0^T\int_{\mathbb{R}^d}\frac{\Theta(|\boldsymbol{b}_t^h|)}{1+|x|}\,d\mu_t^h dt \le \int_0^T\int_{\mathbb{R}^d}\frac{\Theta(|\boldsymbol{b}_t|)}{1+|x|}\,d\mu_t dt < +\infty \tag{18}$$

for some strictly convex function $\Theta:\mathbb{R}^+\to\mathbb{R}$ *having a more than linear growth at infinity;*

(iv) the PDE fulfils the comparison principle in $\mathscr{L}_{\boldsymbol{b}}$*.*

Then $\mu_t = \boldsymbol{X}(t,\cdot)_\#\bar\mu$ *and* $x\mapsto\boldsymbol{X}^h(\cdot,x)$ *converge to* $x\mapsto\boldsymbol{X}(\cdot,x)$ *in* $\bar\mu$*-measure, i.e.*

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

Proof. Following the same strategy used in the proof of the superposition principle, we push $\bar\mu$ onto the graph of the map $x\mapsto\boldsymbol{X}^h(\cdot,x)$, i.e.

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

and we obtain, using (18) and the same argument used in Step 2 of the proof of the superposition principle, that $\boldsymbol{\eta}^h$ is tight in $\mathscr{M}_+(\mathbb{R}^d\times\Gamma_T)$.

Let now $\boldsymbol{\eta}$ be any limit point of $\boldsymbol{\eta}^h$. Using the same argument used in Step 3 of the proof of the superposition principle and (18) we obtain that $\boldsymbol{\eta}$ is concentrated on pairs (x,γ) with γ absolutely continuous solution of the ODE relative to $\boldsymbol{b}$ starting from x. Indeed, this argument was using only the property

$$\lim_{h\to\infty}\int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_t^h-\boldsymbol{c}_t|}{1+|x|}\,d\mu_t^h\,dt=\int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_t-\boldsymbol{c}_t|}{1+|x|}\,d\mu_t\,dt$$

for any continuous function $\boldsymbol{c}$ with compact support in $(0,T)\times\mathbb{R}^d$, and this property is ensured by Lemma 23 below.
Let $\mu_t:=(\boldsymbol{e}_t)_\#\boldsymbol{\eta}$ and notice that $\mu_t^h=(\boldsymbol{e}_t)_\#\boldsymbol{\eta}^h$, hence $\mu_t^h\to\mu_t$ narrowly for any $t\in[0,T]$. As $\mu_t^h\in\mathscr{L}_{\boldsymbol{b}^h}$, assumption (i) gives that $\mu_t\in\mathscr{L}_{\boldsymbol{b}}$ and assumption (iv) together with Theorem 18 imply that $\boldsymbol{\eta}$ is concentrated on the graph of the map $x\mapsto\boldsymbol{X}(\cdot,x)$, where $\boldsymbol{X}$ is the unique $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow. We have thus obtained that

$$\big(x,\boldsymbol{X}^h(\cdot,x)\big)_\#\bar\mu\quad\rightharpoonup\quad\big(x,\boldsymbol{X}(\cdot,x)\big)_\#\bar\mu.$$

By applying the following general principle we conclude. □

Lemma 22 (Narrow convergence and convergence in measure) *Let $v_h,\,v:X\to Y$ be Borel maps and let $\bar\mu\in\mathscr{M}_+(X)$. Then $v_h\to v$ in $\bar\mu$-measure iff*

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

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

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

and notice that

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

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

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

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

and since ϵ is arbitrary the proof is achieved. □

Lemma 23 *Let $A \subset \mathbb{R}^m$ be an open set, and let $\sigma^h \in \mathscr{M}_+(A)$ be narrowly converging to $\sigma \in \mathscr{M}_+(A)$. Let $\boldsymbol{f}^h \in L^1(A, \sigma^h, \mathbb{R}^k)$, $\boldsymbol{f} \in L^1(A, \sigma, \mathbb{R}^k)$ and assume that*

(i) $\boldsymbol{f}^h \sigma^h$ weakly converge, in the duality with $C_c(A; \mathbb{R}^k)$, to $\boldsymbol{f}\sigma$;

(ii) $\limsup_{h\to\infty} \int_A \Theta(|\boldsymbol{f}^h|)\, d\sigma^h \le \int_A \Theta(|\boldsymbol{f}|)\, d\sigma < +\infty$ for some strictly convex function $\Theta : \mathbb{R}^+ \to \mathbb{R}$ having a more than linear growth at infinity.

Then $\int_A |\boldsymbol{f}^h - \boldsymbol{c}|\, d\sigma^h \to \int_A |\boldsymbol{f} - \boldsymbol{c}|\, d\sigma$ for any $\boldsymbol{c} \in C_b(A; \mathbb{R}^k)$.

Proof. We consider the measures $\nu^h := (x, \boldsymbol{f}^h(x))_\# \sigma^h$ in $A \times \mathbb{R}^k$ and we assume, possibly extracting a subsequence, that $\nu^h \rightharpoonup \nu$, with $\nu \in \mathscr{M}_+(A \times \mathbb{R}^k)$, in the duality with $C_c(A \times \mathbb{R}^k)$. Using condition (ii), the narrow convergence of σ^h and a truncation argument it is easy to see that the convergence actually occurs for any continuous test function $\psi(x, y)$ satisfying

$$\lim_{|y|\to\infty} \frac{\sup_x |\psi(x,y)|}{\Theta(|y|)} = 0.$$

Furthermore, for nonnegative continuous functions ψ, we have also

$$\int_{A\times\mathbb{R}^k} \psi\, d\nu \le \liminf_{h\to\infty} \int_{A\times\mathbb{R}^k} \psi\, d\nu_h. \tag{19}$$

Then, choosing test functions $\psi = \psi(x) \in C_b(A)$, the convergence of σ^h to σ gives

$$\int_{A\times\mathbb{R}^k} \psi\, d\nu = \int_A \psi\, d\sigma$$

and therefore, according to the disintegration theorem, we can represent ν as

$$\int_{A\times\mathbb{R}^k} \psi(x,y)\, d\nu(x,y) = \int_A \left(\int_{\mathbb{R}^k} \psi(x,y)\, d\nu_x(y) \right) d\sigma(x) \tag{20}$$

for a suitable Borel family of probability measures ν_x in $\mathbb{R}^k$. Next, we can use $\psi(x)y_j$ as test functions and assumption (i), to obtain

$$\lim_{h\to\infty} \int_A \boldsymbol{f}^h_j \psi\, d\mu^h = \lim_{h\to\infty} \int_{A\times\mathbb{R}^k} \psi(x) y_j\, d\nu^h = \int_A \psi(x) \left(\int_{\mathbb{R}^k} y_j\, d\nu_x(y) \right) d\sigma(x).$$

As ψ and j are arbitrary, this means that the first moment ν_x, *i.e.* $\int y\, d\nu_x$, is equal to $\boldsymbol{f}(x)$ for σ-a.e. x.
On the other hand, choosing $\psi(y) = \Theta(|y|)$ as test function in (19), assumption (ii) gives

$$\int_A \int_{\mathbb{R}^k} \Theta(|y|)\, d\nu_x(y)\, d\sigma(x) \le \liminf_{h\to\infty} \int_{A\times\mathbb{R}^k} \Theta(|y|)\, d\nu^h = \limsup_{h\to\infty} \int_A \Theta(|\boldsymbol{f}^h|)\, d\sigma^h = \int_A \Theta(|\boldsymbol{f}|)\, d\sigma,$$

hence $\int \Theta(|y|)\, d\nu_x = \boldsymbol{f}(x) = \Theta\left(\left|\int y\, d\nu_x\right|\right)$ for σ-a.e. x. As Θ is strictly convex, this can happen only if $\nu_x = \delta_{\boldsymbol{f}(x)}$ for σ-a.e. x.
Finally, taking into account the representation (20) of ν with $\nu_x = \delta_{\boldsymbol{f}(x)}$, the convergence statement can be achieved just choosing the test function $\psi(x, y) = |y - \boldsymbol{c}(x)|$. □

4 Vector fields with a Sobolev spatial regularity

Here we discuss the well-posedness of the continuity or transport equations assuming the $\boldsymbol{b}_t(\cdot)$ has a Sobolev regularity, following [53]. Then, the general theory previously developed provides existence, uniqueness and stability of the $\mathscr{L}$-Lagrangian flow, with $\mathscr{L} := L^\infty(L^1) \cap L^\infty(L^\infty)$. We denote by $I \subset \mathbb{R}$ an open interval.

Definition 24 (Renormalized solutions) *Let $\boldsymbol{b} \in L^1_{\rm loc}\left(I; L^1_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ be such that $D\cdot\boldsymbol{b}_t = \operatorname{div}\boldsymbol{b}_t\mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t\in I$, with*

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

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

$$c := \frac{d}{dt}w + \boldsymbol{b}\cdot\nabla w \in L^1_{\rm loc}(I\times\mathbb{R}^d). \tag{21}$$

Then, we say that w is a renormalized solution of (21) if

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

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

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

Notice also that the concept makes sense, choosing properly the class of "test" functions β, also for w that do not satisfy (21), or are not even locally integrable. This is particularly relevant in connection with DiPerna-Lions's existence theorem for Boltzmann equation , or with the case when w is the characteristic of an unbounded vector field $\boldsymbol{b}$.
This concept is also reminiscent of Kruzkhov's concept of *entropy* solution for a scalar conservation law

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

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

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

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

Remark 25 (Time continuity) Using the fact that both $t\mapsto w_t$ and $t\mapsto\beta(w_t)$ have a uniformly continuous representative (w.r.t. the $w^* - L^\infty_{\rm loc}$ topology), we obtain that, for any renormalized solution w, $t\mapsto w_t$ has a unique representative which is continuous w.r.t. the $L^1_{\rm loc}$ topology. The proof follows by a classical weak-strong convergence argument:

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

provided β is *strictly* convex. In the case of scalar conservation laws there are analogous results [82], [73].

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

$$\begin{aligned}\mathscr{L} &:= \Big\{w \in L^\infty\Big([0,T];L^1(\mathbb{R}^d)\Big) \cap L^\infty\Big([0,T];L^\infty(\mathbb{R}^d)\Big) : \\ &\quad w \in C\Big([0,T];w^* - L^\infty(\mathbb{R}^d)\Big)\Big\}.\end{aligned} \tag{22}$$

Theorem 26 (Comparison principle) *Assume that*

$$\frac{|\boldsymbol{b}|}{1+|x|} \in L^1\Big([0,T];L^\infty(\mathbb{R}^d)\Big) + L^1\Big([0,T];L^1(\mathbb{R}^d)\Big), \tag{23}$$

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

$$[\operatorname{div} \boldsymbol{b}_t]^- \in L^1_{\rm loc}\Big([0,T) \times \mathbb{R}^d\Big). \tag{24}$$

Setting $\boldsymbol{b}_t \equiv 0$ *for* $t < 0$*, assume in addition that* any *solution of* (21) *in* $(-\infty,T) \times \mathbb{R}^d$ *is renormalized. Then the comparison principle for the continuity equation holds in the class* $\mathscr{L}$*.*

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

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

Notice that

$$\beta_\epsilon(t) \uparrow t^+ \quad \text{as } \epsilon \downarrow 0, \qquad t\beta'_\epsilon(t) - \beta_\epsilon(t) \in [0,\epsilon]. \tag{25}$$

We know that

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

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

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

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

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

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

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

The second integral can be estimated with

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

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

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

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

Remark 27 It would be nice to have a completely non-linear comparison principle between renormalized solutions, as in the Kruzkhov theory. Here, on the other hand, we rather used the fact that the difference of the two solutions is renormalized.

In any case, Di Perna and Lions proved that all distributional solutions are renormalized when there is a Sobolev regularity with respect to the spatial variables.

Theorem 28 *Let $b \in L^1_{\mathrm{loc}}\left(I; W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right)$ and let $w \in L^\infty_{\mathrm{loc}}(I \times \mathbb{R}^d)$ be a distributional solution of* (21)*. Then w is a renormalized solution.*

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

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

to obtain

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

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

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

When we let $\epsilon \downarrow 0$ the convergence in the distribution sense of all terms in the identity above is trivial, with the exception of the last one. To ensure its convergence to zero, it seems necessary to show that $r^\epsilon \to 0$ strongly in L^1_{loc} (remember that $\beta'(w^\epsilon)$ is locally equibounded w.r.t. ϵ). This is indeed the case, and it is exactly here that the Sobolev regularity plays a role. □

Proposition 29 (Strong convergence of commutators) *If $w \in L^\infty_{\mathrm{loc}}(I \times \mathbb{R}^d)$ and $\boldsymbol{b} \in L^1_{\mathrm{loc}}\left(I; W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right)$ we have*

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

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

$$r^\epsilon(t,x)=\int_{\mathbb{R}^d} w(t,x-\epsilon y)\frac{(\boldsymbol{b}_t(x-\epsilon y)-\boldsymbol{b}_t(x))\cdot\nabla\rho(y)}{\epsilon}\,dy-(w\mathrm{div}\,\boldsymbol{b}_t)*\rho_\epsilon(x). \tag{26}$$

Introducing the commutators in the (easier) conservative form

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

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

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

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

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

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

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

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

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

The elementary identity

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

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

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

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

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

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

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

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

Theorem 30 (Existence and uniqueness of $\mathscr{L}$-Lagrangian flows) *Let* $\boldsymbol{b} \in L^1\left([0,T]; W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ *be satisfying*

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

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

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

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

$$\frac{d}{dt}w + D_x \cdot (\boldsymbol{b}w) = 0 \tag{27}$$

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

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

Therefore, any weak limit of w^h solves (27). $\square$

Notice also that, choosing for instance a Gaussian, we obtain that the $\mathscr{L}$-Lagrangian flow is well defined up to $\mathscr{L}^d$-negligible sets (and independent of $\bar{\mu} \ll \mathscr{L}^d$, thanks to Remark 17).
It is interesting to compare our characterization of Lagrangian flows with the one given in [53]. Heuristically, while the Di Perna-Lions one is based on the semigroup of transformations $x \mapsto \boldsymbol{X}(t,x)$, our one is based on the properties of the map $x \mapsto \boldsymbol{X}(\cdot,x)$.

Remark 31 The definition of the flow in [53] is based on the following three properties:

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

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

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

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

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

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

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

It is immediate to check that $\tilde{\boldsymbol{X}}(\cdot,x)$ is an integral solution of the ODE in $[s',T]$ for $\mathscr{L}^d$-a.e. x and that $\tilde{\boldsymbol{X}}(t,\cdot)_\#\bar\mu$ is bounded by $C^2\mathscr{L}^d$. Then, Theorem 30 (with s' as initial time) gives $\tilde{\boldsymbol{X}}(\cdot,x) = \boldsymbol{X}(\cdot,s',x)$ in $[s',T]$ for $\mathscr{L}^d$-a.e. x, whence (c) follows.

Moreover, the stability Theorem 21 can be read in this context as follows. We state it for simplicity only in the case of equi-bounded vectorfields (see [9] for more general results).

Theorem 32 (Stability) *Let $\boldsymbol{b}^h$, $\boldsymbol{b}\in L^1\left([0,T];W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$, let $\boldsymbol{X}^h$, $\boldsymbol{X}$ be the $\mathscr{L}$-Lagrangian flows relative to $\boldsymbol{b}^h$, $\boldsymbol{b}$, let $\bar\mu=\rho\mathscr{L}^d\in\mathscr{M}_+(\mathbb{R}^d)$ and assume that*

(i) $\boldsymbol{b}^h\to\boldsymbol{b}$ *in* $L^1_{\rm loc}\left((0,T)\times\mathbb{R}^d\right)$*;*

(ii) $|\boldsymbol{b}_h|\le C$ *for some constant C independent of h;*

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

Then,

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

Proof. It is not restrictive, by an approximation argument, to assume that ρ has a compact support. Under this assumption, (i) and (iii) ensure that $\mu^h_t\le M\chi_{B_R}\mathscr{L}^d$ for some constants M and R independent of h and t. Denoting by μ_t the weak limit of μ_t^h, choosing $\Theta(z)=|z|^2$ in (iii) of Theorem 21, we have to check that

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

Let $\epsilon > 0$ and let $B \subset (0,T) \times B_R$ be an open set given by Egorov theorem, such that $\boldsymbol{b}_h \to \boldsymbol{b}$ uniformly on $[0,T] \times B_R \setminus B$ and $\mathscr{L}^{d+1}(B) < \epsilon$. Let also $\tilde{\boldsymbol{b}}_\epsilon$ be such that $|\tilde{\boldsymbol{b}}_\epsilon| \le C$ and $\tilde{\boldsymbol{b}}_\epsilon = \boldsymbol{b}$ on $[0,T] \times B_R \setminus B$. We write

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_h|^2}{1+|x|}\, d\mu_t^h dt - \int_0^T \int_{\mathbb{R}^d} \frac{|\tilde{\boldsymbol{b}}_\epsilon|^2}{1+|x|}\, d\mu_t^h dt = \int_{[0,T]\times B_R \setminus B} \frac{|\boldsymbol{b}_h|^2 - |\tilde{\boldsymbol{b}}_\epsilon|^2}{1+|x|}\, d\mu_t^h dt + \int_B \frac{|\boldsymbol{b}_h|^2 - |\tilde{\boldsymbol{b}}|^2}{1+|x|}\, d\mu_t^h dt,$$

so that

$$\limsup_{h\to\infty} \left| \int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_h|^2}{1+|x|}\, d\mu_t^h dt - \int_0^T \int_{\mathbb{R}^d} \frac{|\tilde{\boldsymbol{b}}_\epsilon|^2}{1+|x|}\, d\mu_t dt \right| \le 2C^2 M \epsilon.$$

As ϵ is arbitrary and

$$\lim_{\epsilon\to 0} \int_0^T \int_{\mathbb{R}^d} \frac{|\tilde{\boldsymbol{b}}_\epsilon|^2}{1+|x|}\, d\mu_t dt = \int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}|^2}{1+|x|}\, d\mu_t dt$$

this proves that (28) is fulfilled. □

Finally, we conclude this section with the illustration of some recent results [64], [13], [14] that seem to be more specific of the Sobolev case, concerned with the "differentiability" w.r.t. to x of the flow $\boldsymbol{X}(t,x)$. These results provide a sort of bridge with the standard Cauchy-Lipschitz calculus:

Theorem 33 *There exist Borel maps $L_t : \mathbb{R}^d \to M^{d\times d}$ satisfying*

$$\lim_{h\to 0} \frac{\boldsymbol{X}(t,x+h) - \boldsymbol{X}(t,x) - L_t(x)h}{|h|} = 0 \quad \textit{locally in measure}$$

for any $t \in [0,T]$. If, in addition, we assume that

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

then the flow has the following "local" Lipschitz property: for any $\epsilon > 0$ there exists a Borel set A with $\bar{\mu}(\mathbb{R}^d \setminus A) < \epsilon$ such that $\boldsymbol{X}(t,\cdot)|_A$ is Lipschitz for any $t \in [0,T]$.

According to this result, L can be thought as a (very) weak derivative of the flow $\boldsymbol{X}$. It is still not clear whether the local Lipschitz property holds in the $W^{1,1}_{\mathrm{loc}}$ case, or in the BV_{loc} case discussed in the next section.

5 Vector fields with a BV spatial regularity

In this section we prove the renormalization Theorem 28 under the weaker assumption of a BV dependence w.r.t. the spatial variables, but still assuming that

$$D \cdot \boldsymbol{b}_t \ll \mathscr{L}^d \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{29}$$

Theorem 34 *Let $\boldsymbol{b} \in L^1_{\rm loc}\big((0,T);BV_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\big)$ be satisfying (29). Then any distributional solution $w \in L^\infty_{\rm loc}\big((0,T)\times\mathbb{R}^d\big)$ of*

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

is a renormalized solution.

We try to give reasonably detailed proof of this result, referring to the original paper [7] for minor details. Before doing that we set up some notation, denoting by $D\boldsymbol{b}_t = D^a\boldsymbol{b}_t + D^s\boldsymbol{b}_t = \nabla\boldsymbol{b}_t\mathscr{L}^d + D^s\boldsymbol{b}_t$ the Radon–Nikodym decomposition of $D\boldsymbol{b}_t$ in absolutely continuous and singular part w.r.t. $\mathscr{L}^d$. We also introduce the measures $|D\boldsymbol{b}|$ and $|D^s\boldsymbol{b}|$ by integration w.r.t. the time variable, i.e.

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

We shall also assume, by the locality of the arguments involved, that $\|w\|_\infty \le 1$.
We are going to find two estimates on the commutators, quite sensitive to the choice of the convolution kernel, and then combine them in a (pointwise) kernel optimization argument.
Step 1 (anisotropic estimate). Let us start from the expression

$$r^\epsilon(t,x) = \int_{\mathbb{R}^d} w(t,x-\epsilon y)\frac{(\boldsymbol{b}_t(x-\epsilon y)-\boldsymbol{b}_t(x))\cdot\nabla\rho(y)}{\epsilon}\,dy - (w\mathrm{div}\,\boldsymbol{b}_t)*\rho_\epsilon(x) \tag{30}$$

of the commutators $(\boldsymbol{b}\cdot\nabla w)*\rho_\epsilon - \boldsymbol{b}\cdot(\nabla(w*\rho_\epsilon))$: since $\boldsymbol{b}_t \notin W^{1,1}$ we cannot use anymore the strong convergence of the difference quotients. However, for any function $u \in BV_{\rm loc}$ and any $z\in\mathbb{R}^d$ with $|z|<\epsilon$ we have a classical L^1 estimate on the difference quotients

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

where $Du = (D_1u,\ldots,D_du)$ stands for the distributional derivative of u, $D_zu = \langle Du,z\rangle = \sum_i z_iD_iu$ denotes the component along z of Du and K_ϵ is the open ϵ-neighbourhood of K. Its proof follows from an elementary smoothing and lower semicontinuity argument.
We notice that, setting $Db_t = M_t|Db_t|$, we have

$$D_z\langle b_t,\nabla\rho(z)\rangle = \langle M_t(\cdot)z,\nabla\rho(z)\rangle|Db| \qquad \forall z\in\mathbb{R}^d$$

and therefore the L^1 estimate on difference quotients gives the *anisotropic estimate*

$$\limsup_{\epsilon\downarrow 0}\int_K |r^\epsilon|\,dx \le \int_K\int_{\mathbb{R}^d}|\langle M_t(x)z,\nabla\rho(z)\rangle|\,dzd|Db|(t,x) + d|D^ab|(K) \tag{31}$$

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

Step 2 (isotropic estimate). 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 Jensen's inequality 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

$$\int_K |\hat\mu_\epsilon|\,dt \le |\mu|(K_\epsilon) \quad \text{for any compact set } K \subset \mathbb{R}, \tag{32}$$

where K_ϵ is again the open ϵ neighbourhood of K. A density argument based on (32) 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

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

for $\mathscr{L}^1$-a.e. x (the exceptional set possibly depends on ϵ). In this way we have canonically split the difference quotient of u 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 $|D^s u|(K)$.
If we fix the direction z of the difference quotient, the slicing theory of BV functions gives that this decomposition can be carried on also in d dimensions, showing that the difference quotients

$$\frac{\boldsymbol{b}_t(x+\epsilon z) - \boldsymbol{b}_t(x)}{\epsilon}$$

can be canonically split into two parts, the first one strongly converging in $L^1_{\rm loc}(\mathbb{R}^d)$ to $\nabla \boldsymbol{b}_t(x) z$, and the second one having an L^1 norm on K asymptotically smaller than $|\langle D^s \boldsymbol{b}_t, z\rangle|(K)$. 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 the *isotropic estimate*

$$\limsup_{\epsilon\downarrow 0} \int_K |r^\epsilon|\,dx \le \left(\int_K \int_{\mathbb{R}^d} |z||\nabla\rho(z)|\,dz\right)\,d|D^s \boldsymbol{b}|(t,x) \tag{33}$$

for any compact set $K \subset (0,T)\times\mathbb{R}^d$.
Step 3 (reduction to a pointwise optimization problem). Roughly speaking, the isotropic estimate is useful in the regions where the absolutely continuous part is the dominant one, so that $|D^s b|(K) << |D^a b|(K)$, while the anisotropic one turns out to be useful in the regions where the dominant part is the singular one, i.e. $|D^a b|(K) << |D^s b|(K)$. Since the two measures are mutually singular, for a typical small ball K only one of these two situations occurs. Let us see how the two estimates can be combined: coming back to the smoothing scheme, we have

$$\frac{d}{dt}\beta(w^\epsilon) + \boldsymbol{b}\cdot\nabla\beta(w^\epsilon) - \beta'(w^\epsilon)c * \rho_\epsilon = \beta'(w^\epsilon) r^\epsilon. \tag{34}$$

Let L be the supremum of $|\beta'|$ on $[-1,1]$. Then, since K is an arbitrary compact set, (33) tells us that any limit measure ν of $|\beta'(w^\epsilon)r^\epsilon|\mathscr{L}^d$ as $\epsilon \downarrow 0$ satisfies

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

and, in particular, is singular with respect to $\mathscr{L}^d$. On the other hand, the estimate (31) tells also us that

$$\nu \le L \int_{\mathbb{R}^d} |\langle M.(\cdot)z, \nabla\rho(z)\rangle|\, dz |D\boldsymbol{b}| + d|D^a b|(K).$$

The second estimate and the singularity of ν with respect to $\mathscr{L}^d$ give

$$\nu \le L \int_{\mathbb{R}^d} |\langle M.(\cdot)z, \nabla\rho(z)\rangle|\, dz |D^s\boldsymbol{b}|. \tag{35}$$

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 (34) tells us that the "defect" measure

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

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

$$|\sigma| \le L\Lambda(M.(\cdot), \rho)|D^s\boldsymbol{b}| \quad \text{with} \quad \Lambda(N,\rho) := \int_{\mathbb{R}^d} |\langle Nz, \nabla\rho(z)\rangle|\, dz. \tag{36}$$

For (x,t) *fixed*, we are thus led to the minimum problem

$$G(N) := \inf\left\{\Lambda(N,\rho) :\ \rho \in C^\infty_c(B_1),\ \rho \ge 0,\ \int_{\mathbb{R}^d} \rho = 1\right\} \tag{37}$$

with $N = M_t(x)$. Indeed, notice that (36) gives

$$|\sigma| \le L \inf_{\rho\in D} \Lambda(M.(\cdot),\rho)|D^s\boldsymbol{b}|$$

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

$$|\sigma| \le LG(M.(\cdot))|D^s\boldsymbol{b}|. \tag{38}$$

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

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

Hence, recalling the definition of $|D^s\boldsymbol{b}|$, the trace of $M_t(x)$ vanishes for $|D^s\boldsymbol{b}|$-a.e. (t,x). Applying the following lemma, a courtesy of Alberti, and using (38) we obtain that $\sigma = 0$, thus concluding the proof.

Lemma 35 (Alberti) *For any $d \times d$ matrix N the infimum in* (37) *is* $|\mathrm{trace}\, N|$.

Proof. We have to build kernels ρ in such a way that the field Nz is as much tangential as possible to the level sets of ρ. Notice first that the lower bound follows immediately by the identity

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

Hence, we have to show only the upper bound. Again, by the identity

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

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

$$\int_{\mathbb{R}^d} |\mathrm{div}\,(Nz\rho(z))|\, dz \le \frac{2}{T}. \tag{39}$$

The heuristic idea is (again...) 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 convolution kernel θ with compact support, it turns out that the function

$$\rho(z) := \frac{1}{T}\int_0^T \theta(e^{-tN}z)e^{-t\,\mathrm{trace}\, N}\, dt \tag{40}$$

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 $\varphi \in C^\infty_c(\mathbb{R}^d)$ we have

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

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

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

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 $\varphi \in C^\infty_c(\mathbb{R}^d)$, so that $TD_z \cdot (Nz\mu_x) = \delta_x - \delta_{e^{TN}x}$. □

The original argument in [7] was slightly different and used, instead of Lemma 35, a much deeper result, still due to Alberti, saying that for a $BV_{\rm loc}$ function $u:\mathbb{R}^d\to\mathbb{R}^m$ the matrix $M(x)$ in the polar decomposition $Du=M|Du|$ has rank 1 for $|D^s u|$-a.e. x, i.e. there exist unit vectors $\xi(x)\in\mathbb{R}^d$ and $\eta(x)\in\mathbb{R}^m$ such that $M(x)z=\eta(x)\langle z,\xi(x)\rangle$. In this case the asymptotically optimal kernels are much easier to build, by mollifying in the ξ direction much faster than in all other ones. This is precisely what Bouchut and Lions did in some particular cases (respectively "Hamiltonian" vector fields and piecewise Sobolev ones).
As in the Sobolev case we can now obtain from the general theory given in Section 3 *existence and uniqueness* of $\mathscr{L}$-Lagrangian flows, with $\mathscr{L}=L^\infty(L^1)\cap L^\infty(L^\infty)$: we just replace in the statement of Theorem 30 the assumption $\boldsymbol{b}\in L^1\left([0,T];W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ with $\boldsymbol{b}\in L^1\left([0,T];BV_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$, assuming as usual that $D\cdot\boldsymbol{b}_t\ll\mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t\in[0,T]$.
Analogously, with the same replacements in Theorem 32 (for $\boldsymbol{b}$ and $\boldsymbol{b}^h$) we obtain *stability* of $\mathscr{L}$-Lagrangian flows.

6 Applications

6.1. A system of conservation laws. Let us consider the Cauchy problem (studied in one space dimension by Keyfitz–Kranzer in [63])

$$\frac{d}{dt}u+\sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left(\boldsymbol{f}_i(|u|)u\right)=0,\qquad u:\mathbb{R}^d\times(0,+\infty)\to\mathbb{R}^k \tag{41}$$

with the initial condition $u(\cdot,0)=\bar u$. Here $\boldsymbol{f}:\mathbb{R}\to\mathbb{R}^d$ is a C^1 function.
In a recent paper [32] Bressan showed that the problem can be ill-posed for L^∞ initial data and he conjectured that it could be well posed for BV initial data, suggesting to extend to this case the classical method of characteristics. In [8] we proved that this procedure can really be implemented, thanks to the results in [7], for initial data $\bar u$ such that $\bar\rho:=|\bar u|\in BV\cap L^\infty$, with $1/|\bar u|\in L^\infty$. Later on, in a joint work with Bouchut and De Lellis [10], we proved that the lower bound on $\bar\rho$ is not necessary and, moreover, we proved that the solution built in [8] is unique in a suitable class of admissible functions: those whose modulus ρ satisfies the scalar PDE

$$\frac{d}{dt}\rho+\sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left(\boldsymbol{f}_i(\rho)\rho\right)=0 \tag{42}$$

in the Kruzhkov sense (i.e. $\eta(\rho)_t+D_x\cdot(\boldsymbol{q}(\rho))\le 0$ for any convex entropy-entropy flux pair $(\eta,\boldsymbol{q})$, here $(s\boldsymbol{f})'(s)\eta'(s)=\boldsymbol{q}'(s)$), with the initial condition $\rho(0,\cdot)=\bar\rho$.
Notice that the regularity theory for this class of solutions gives that $\rho\in L^\infty\cap BV_{\rm loc}\left([0,+\infty)\times\mathbb{R}^d\right)$, due to the BV regularity and the boundedness of $|\bar u|$. Furthermore the maximum principle gives $0<1/\rho\le 1/|\bar u|\in L^\infty$.
In order to obtain the (or, better, a) solution u we can *formally* decouple the system, writing

$$u=\theta\rho,\qquad \bar u=\bar\theta\bar\rho,\qquad |\theta|=|\bar\theta|=1,$$

thus reducing the problem to the system (decoupled, if one neglects the constraint $|\theta| = 1$) of transport equations

$$\theta_t + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} (\boldsymbol{f}_i(\rho)\theta) = 0 \tag{43}$$

with the initial condition $\theta(0, \cdot) = \bar{\theta}$.
A *formal* solution of the system, satisfying also the constraint $|\theta| = 1$, is given by

$$\theta(t, x) := \bar{\theta}\left([\boldsymbol{X}(t, \cdot)]^{-1}(x)\right),$$

where $\boldsymbol{X}(t, \cdot)$ is the flow associated to $\boldsymbol{f}(\rho)$. Notice that the non-autonomous vector field $\boldsymbol{f}(\rho)$ is bounded and of class BV_{loc}, but the theory illustrated in these lectures is not immediately applicable because its divergence *is not* absolutely continuous with respect to $\mathscr{L}^{d+1}$. In this case, however, a simple argument still allows the use of the theory, representing $\boldsymbol{f}(\rho)$ as a part of the autonomous vector field $\boldsymbol{b} := (\rho, \rho\boldsymbol{f}(\rho))$ in $\mathbb{R}^+ \times \mathbb{R}^d$. This new vector field is still BV_{loc} and bounded, and it is divergence-free due to (42).
At this point, it is not hard to see that the reparameterization of the flow $(t(s), \boldsymbol{x}(s))$ associated to $\boldsymbol{b}$

$$\left(\dot{t}(s), \dot{\boldsymbol{x}}(s)\right) = (\rho(t(s), \boldsymbol{x}(s)), \boldsymbol{f}(\rho(t(s), \boldsymbol{x}(s)))\rho(t(s), \boldsymbol{x}(s)))$$

defined by $\tilde{\boldsymbol{x}}(t) = \boldsymbol{x}(t(s)^{-1}(t))$ (and here we use the assumption $\rho > 0$) defines a flow for the vector field $\boldsymbol{f}(\rho)$ we were originally interested to.
In this way we get a kind of formal, or pointwise, solution of the system (42), that could indeed be very far from being a *distributional* solution.
But here comes into play the stability theorem, showing that all formal computations above can be justified just assuming first $(\rho, \boldsymbol{f}(\rho))$ smooth, and then by approximation (see [8] for details).

6.2. Lagrangian solutions of semi-geostrophic equations. The semigeostrophic equations are a simplifies model of the atmosphere/ocean flows [45], described by the system of transport equations

$$\text{(SGE)} \qquad \begin{cases} \dfrac{d}{dt}\partial_2 p + \boldsymbol{u} \cdot \nabla \partial_2 p = -\boldsymbol{u}_2 + \partial_1 p \\ \dfrac{d}{dt}\partial_1 p + \boldsymbol{u} \cdot \nabla \partial_1 p = -\boldsymbol{u}_1 - \partial_2 p \\ \dfrac{d}{dt}\partial_3 p + \boldsymbol{u} \cdot \nabla \partial_3 p = 0. \end{cases}$$

Here $\boldsymbol{u}$, the velocity, is a divergence-free field, p is the pressure and $\rho := -\partial_3 p$ represents the density of the fluid. We consider the problem in $[0, T] \times \Omega$, with Ω bounded and convex. Initial conditions are given on the pressure and a no-flux condition through $\partial\Omega$ is imposed for all times. Introducing the modified pressure $P_t(x) := p_t(x) + (x_1^2 + x_2^2)/2$, (SGE) can be written in a more compact form as

$$\frac{d}{dt}\nabla P + \boldsymbol{u} \cdot \nabla^2 P = J(\nabla P - x) \quad \text{with} \quad J := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{44}$$

Existence (and uniqueness) of solutions are still open for this problem. In [20] and [46], existence results have been obtained in the so-called dual coordinates, where we replace the physical variable x by $X = \nabla P_t(x)$. Under this change of variables, *and assuming P_t to be convex*, the system becomes

$$\frac{d}{dt}\alpha_t + D_x \cdot (\boldsymbol{U}_t \alpha_t) = 0 \quad \text{with} \quad \boldsymbol{U}_t(X) := J\left(X - \nabla P_t^*(X)\right) \tag{45}$$

with $\alpha_t := (\nabla P_t)_\#(\mathscr{L}_\Omega)$ (here we denote by $\mathscr{L}_\Omega$ the restriction of $\mathscr{L}^d$ to Ω). Indeed, for any test function φ we can use the fact that $\boldsymbol{u}$ is divergence-free to obtain:

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^d} \varphi\, d\alpha_t &= \int_{\mathbb{R}^d} \nabla\varphi(\nabla P_t) \cdot \frac{d}{dt}\nabla P_t\, dx \\
&= \int_{\mathbb{R}^d} \nabla\varphi(\nabla P_t) \cdot J(\nabla P_t - x)\, dx + \int_{\mathbb{R}^d} \nabla\varphi(\nabla P_t)\nabla^2 P_t \cdot \boldsymbol{u}\, dx \\
&= \int_{\mathbb{R}^d} \nabla\varphi \cdot J(X - \nabla P_t^*)\, d\alpha_t + \int_{\mathbb{R}^d} \nabla(\varphi \circ \nabla P_t) \cdot \boldsymbol{u}\, dx \\
&= \int_{\mathbb{R}^d} \nabla\varphi \cdot \boldsymbol{U}_t\, d\alpha_t.
\end{aligned}$$

Existence of a solution to (45) can be obtained by a suitable time discretization scheme. Now the question is: can we go back to the original physical variables ? An important step forward has been achieved by Cullen and Feldman in [47], with the concept of *Lagrangian* solution of (SGE).
Taking into account that the vector field $\boldsymbol{U}_t(X) = J(X - \nabla P_t^*(X))$ is BV, bounded and divergence-free, there is a well defined, stable and measure preserving flow $\boldsymbol{X}(t, X) = \boldsymbol{X}_t(X)$ relative to $\boldsymbol{U}$. This flow can be carried back to the physical space with the transformation

$$F_t(x) := \nabla P_t^* \circ \boldsymbol{X}_t \circ \nabla P_0(x),$$

thus defining maps F_t preserving $\mathscr{L}^d_\Omega$.
Using the stability theorem can also show that $Z_t(x) := \nabla P_t(F_t(x))$ solve, in the distributions sense, the Lagrangian form of (44), i.e.

$$\frac{d}{dt} Z_t(x) = J(Z_t - F_t) \tag{46}$$

This provides us with a sort of weak solution of (44), and it is still an open problem how the Eulerian form could be recovered (see Section 7).

7 Open problems, bibliographical notes, and references

Section 2. The material contained in this section is classical. Good references are [56], Chapter 8 of [12], [29] and [53]. For the proof of the area formula, see for instance [6], [55], [60].
The proof of the second local variant, under the stronger assumption $\int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t|\, d\mu_t dt < +\infty$, is given in Proposition 8.1.8 of [12]. The same proof works under the weaker assumption (6).

Section 3. Many ideas of this section, and in particular the idea of looking at measures in the space of continuous maps to characterize the flow and prove its stability, are borrowed from [7], dealing with BV vector fields. Later on, the arguments have been put in a more general form, independent of the specific class of vector fields under consideration, in [9]. Here we present a more refined version of [9].
The idea of a probabilistic representation is of course classical, and appears in many contexts (particularly for equations of diffusion type); to my knowledge the first reference in the context of conservation laws and fluid mechanics is [24], where a similar approach is proposed for the incompressible Euler equation (see also [25], [26], [27]): in this case the compact (but neither metrizable, nor separable) space $X^{[0,T]}$, with $X \subset \mathbb{R}^d$ compact, has been considered.
This approach is by now a familiar one also in optimal transport theory, where transport maps and transference plans can be thought in a natural way as measures in the space of minimizing geodesics [76], and in the so called irrigation problems, a nice variant of the optimal transport problem [22]. See also [18] for a similar approach within Mather's theory. The Lecture Notes [84] (see also the Appendix of [69]) contain, among several other things, a comprehensive treatment of the topic of measures in the space of action-minimizing curves, including at the same time the optimal transport and the dynamical systems case (this unified treatment was inspired by [21]). Another related reference is [50].
The superposition principle is proved, under the weaker assumption $\int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t|^p \, d\mu_t dt < +\infty$ for some $p > 1$, in Theorem 8.2.1 of [12], see also [70] for the extension to the case $p = 1$ and to the non-homogeneous continuity equation. Very closely related results, relative to the representation of a vector field as the superposition of "elementary" vector fields associated to curves, appear in [77], [18].
In [16] an interesting variant of the stability Theorems 21 and 32 is discussed, peculiar of the case when the limit vector field $\boldsymbol{b}$ is a sufficiently regular gradient. In this case it has been proved in [16] that narrow convergence of μ_t^h to μ_t for all $t \in [0, T]$ and the energy estimate

$$\limsup_{h\to\infty} \int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t^h|^2 \, d\mu_t^h \, dt \le \int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t|^2 \, d\mu_t dt < +\infty$$

are sufficient to obtain the stability property. This is due to the fact that, given μ_t, gradient vector fields minimize $\int_0^T \int |\boldsymbol{c}_t|^2 \, d\mu_t$ among all velocity fields $\boldsymbol{c}_t$ for which the continuity equation $\frac{d}{dt}\mu_t + D_x \cdot (\boldsymbol{c}_t \mu_t) = 0$ holds (see Chapter 8 of [12] for a general proof of this fact, and for references to earlier works of Otto, Benamou-Brenier).
The convergence result in [16] can be used to answer positively a question raised in [59], concerning the convergence of the implicit Euler scheme

$$\boldsymbol{u}_{k+1} \in \text{Argmin} \left[\frac{1}{2h} \int_\Omega |\boldsymbol{u} - \boldsymbol{u}_k|^2 + \int_\Omega F(\nabla \boldsymbol{u}) \, dx \right]$$

(here Ω, Ω' are bounded open in $\mathbb{R}^d$ and $\boldsymbol{u} : \Omega \to \Omega'$) in the case when $F(\nabla \boldsymbol{u})$ depends only, in a convex way, only on the determinant of $\nabla \boldsymbol{u}$. It turns out that, representing as in [59] $\boldsymbol{u}_k$ as the composition of k optimal transport maps, $\boldsymbol{u}_{[t/h]}$ converge as $h \downarrow 0$ to the solution $\boldsymbol{u}_t$ of

$$\frac{d}{dt}\boldsymbol{u}_t = \text{div} \left(\nabla F(\nabla \boldsymbol{u}_t) \right),$$

built in [59] by purely differential methods (coupling a nonlinear diffusion equation for the measures $\beta_t := (\boldsymbol{u}_t)_\#(\mathscr{L}_\Omega)$ in Ω' to a transport equation for $\boldsymbol{u}_t^{-1}$). Existence of solutions (via differential or variational methods) for wider classes of energy densities F is a largely open problem.

Section 4. The definition of renormalized solution and the strong convergence of commutators are entirely borrowed from [53]. See also [54] for the relevance of this concept in connection with the existence theory for Boltzmann equation. The proof of the comparison principle assuming only an $L^1(L^1_{\mathrm{loc}})$ bound (instead of an $L^1(L^\infty)$ one, as in [53], [7]) on the divergence was suggested to me by G.Savaré. The differentiability properties of the flow have been found in [64]: later on, this differentiability property has been characterized and compared with the more classical approximate differentiability [60] in [14], while [13] contains the proof of the stronger "local" Lipschitz properties. Theorem 33 summarizes all these results. The paper [44] contains also more explicit Lipschitz estimates and an independent proof of the compactness of flows. See also [37] for a proof, using radial convolution kernels, of the renormalization property for vector fields satisfying $D_i\boldsymbol{b}^j + D_j\boldsymbol{b}^i \in L^1_{\mathrm{loc}}$.
Both methods, the one illustrated in these notes and the DiPerna–Lions one, are based on abstract compactness arguments and do not provide a rate of convergence in the stability theorem. It would be interesting to find an explicit rate of convergence (in mean with respect to x) of the trajectories. This problem is open even for autonomous, bounded and Sobolev (but not Lipschitz) vector fields.
No general existence result for Sobolev (or even BV) vector fields seems to be known in the infinite-dimensional case: the only reference we are aware of is [23]. Also the investigation of non-Euclidean geometries, *e.g.* Carnot groups and horizontal vector fields, could provide interesting results.
Finally, notice that the theory has a natural invariance, namely if $\boldsymbol{X}$ is a flow relative to $\boldsymbol{b}$, then $\boldsymbol{X}$ is a flow relative to $\tilde{\boldsymbol{b}}$ whenever $\{\tilde{\boldsymbol{b}} \neq \boldsymbol{b}\}$ is $\mathscr{L}^{1+d}$-negligible in $(0,T)\times\mathbb{R}^d$. So a natural question is whether the uniqueness "in the selection sense" might be enforced by choosing a canonical representative $\tilde{\boldsymbol{b}}$ in the equivalence class of $\boldsymbol{b}$: in other words we may think that, for a suitable choice of $\tilde{\boldsymbol{b}}$, the ODE $\dot\gamma(t) = \tilde{\boldsymbol{b}}_t(\gamma(t))$ has a unique absolutely continuous solution starting from x for $\mathscr{L}^d$-a.e. x.

Section 5. Here we followed closely [7]. The main idea of this section, *i.e.* the adaptation of the convolution kernel to the local behaviour of the vector field, has been used at various level of generality in [30], [66], [41] (see also [38], [39] for related results independent of this technique), until the general result [7].
The optimal regularity condition on $\boldsymbol{b}$ ensuring the renormalization property, and therefore the validity of the comparison principle in $\mathscr{L}_{\boldsymbol{b}}$, is still not known. New results, both in the Sobolev and in the BV framework, are presented in [11], [64], [65].
In [15] we investigate in particular the possibility to prove the renormalization property for nearly incompressible $BV_{\mathrm{loc}}\cap L^\infty$ fields $\boldsymbol{b}$: they are defined by the property that there exists a positive function ρ, with $\ln\rho \in L^\infty$, such that the space-time field $(\rho, \rho\boldsymbol{b})$ is divergence free. As in the case of the Keyfitz-Kranzer system, the existence a function ρ with this property seems to be a natural replacement of the condition $D_x\cdot\boldsymbol{b} \in L^\infty$ (and is actually implied by it); as

explained in [10], a proof of the renormalization property in this context would lead to a proof of a conjecture, due to Bressan, on the compactness of flows associated to a sequence of vector fields bounded in $BV_{t,x}$.

Section 6. In connection with the Keyfitz–Kranzer system there are several open questions: in particular one would like to obtain uniqueness (and stability) of the solution in more general classes of admissible functions (partial results in this direction are given in [10]). A strictly related problem is the convergence of the vanishing viscosity method to the solution built in [8]. Also, very little about the regularity of solutions is presently known: we know [49] that BV estimates do not hold and, besides, that the contruction in [8] seems not applicable to more general systems of triangular type, see the counterexample in [43].

In connection with the semi-geostrophic problem, the main problem is the existence of solutions in the physical variables, i.e. in the Eulerian form. A formal argument suggests that, given P_t, the velocity $\boldsymbol{u}$ should be defined by

$$\partial_t \nabla P_t^*(\nabla P_t(x)) + \nabla^2 P_t^*(\nabla P_t(x)) J\big(\nabla P_t(x) - x\big).$$

On the other hand, the a-priori regularity on ∇P_t (ensured by the convexity of P_t) is a BV regularity, and it is still not clear how this formula could be rigorously justified. In this connection, an important intermediate step could be the proof of the $W^{1,1}$ regularity of the maps ∇P_t (see also [33], [34], [35], [36], [80], [81] for the regularity theory of optimal transport maps under regularity assumptions on the initial and final densities).

References

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

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

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

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

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

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

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

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

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

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

[11] L.Ambrosio, G.Crippa & S.Maniglia: *Traces and fine properties of a BD class of vector fields and applications.* Preprint, 2004 (to appear on Annales de Toulouse).

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

[13] L.Ambrosio, M.Lecumberry & S.Maniglia: *Lipschitz regularity and approximate differentiability of the DiPerna-Lions flow.* Preprint, 2005 (available at http://cvgmt.sns.it and to appear on Rend. Sem. Fis. Mat. di Padova).

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

[15] L.Ambrosio, C.De Lellis & J.Malý: *On the chain rule for the divergence of BV like vector fields: applications, partial results, open problems.* Preprint, 2005 (available at http://cvgmt.sns.it).

[16] L.Ambrosio, S.Lisini & G.Savaré: *Stability of flows associated to gradient vector fields and convergence of iterated transport maps.* In preparation.

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

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

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

[20] 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.

[21] P.Bernard & B.Buffoni: *Optimal mass transportation and Mather theory.* Preprint, 2004.

[22] M.Bernot, V.Caselles & J.M.Morel: *Traffic plans.* Preprint, 2004.

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

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

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

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

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

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

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

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

[31] F.Bouchut, F.James & S.Mancini: *Uniqueness and weak stability for multi-dimensional transport equations with one-sided Lipschitz coefficients.* Preprint, 2004 (to appear on Annali Scuola Normale Superiore).

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

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

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

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

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

[37] 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.

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

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

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

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

[42] F.Colombini, T. Luo & J.Rauch: *Uniqueness and nonuniqueness for nonsmooth divergence-free transport.* Preprint, 2003.

[43] G.Crippa & C.De Lellis: *Oscillatory solutions to transport equations.* Preprint, 2005 (available at http://cvgmt.sns.it).

[44] G.Crippa & C.De Lellis: *Estimates for transport equations and regularity of the DiPerna-Lions flow.* In preparation.

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

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

[47] M.Cullen & M.Feldman: *Lagrangian solutions of semigeostrophic equations in physical space.* Preprint, 2003.

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

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

[50] L.De Pascale, M.S.Gelli & L.Granieri: *Minimal measures, one-dimensional currents and the Monge-Kantorovich problem.* Preprint, 2004 (available at http://cvgmt.sns.it).

[51] 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.

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

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

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

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

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

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

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

[59] L.C.Evans, W.Gangbo & O.Savin: *Nonlinear heat flows and diffeomorphisms.* Preprint, 2004.

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

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

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

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

[64] 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.

[65] N.Lerner: *Transport equations with partially BV velocities.* Preprint, 2004.

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

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

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

[69] J.Lott & C.Villani: *Weak curvature conditions and Poincaré inequalities.* Preprint, 2005.

[70] S.Maniglia: *Probabilistic representation and uniqueness results for measure-valued solutions of transport equations.* Preprint, 2005.

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

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

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

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

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

[76] A.Pratelli: *Equivalence between some definitions for the optimal transport problem and for the transport density on manifolds.* preprint, 2003, to appear on Ann. Mat. Pura Appl (available at http://cvgmt.sns.it).

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

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

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

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

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

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

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

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

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