Continuity equations and ODE flows with non-smooth velocity

Luigi Ambrosio[*] and Gianluca Crippa[†]

Edinburgh, April 15-16, 2013

Contents

1 Introduction

In these lectures (based in part on [12, 13], with additional updates and more recent improvements) we study the well-posedness of the Cauchy problem for the homogeneous conservative continuity equation

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

---

[*]Scuola Normale Superiore di Pisa, email: `l.ambrosio@sns.it`

[†]Universität Basel, email: `gianluca.crippa@unibas.ch`

for families time-dependent families μ_t of probability measures in $\mathbb{R}^d$ (or, more generally, signed measures) 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 $\mathbb{R}^d$ and $I \subset \mathbb{R}$ is an interval: 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

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

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 they can stay for an arbitrary time T, then continuing as $x(t) = \frac{1}{4}(t - T - 2c)^2$. Let us consider for instance the Lipschitz approximation (that could easily be made smooth) of $b(\gamma) = \sqrt{|\gamma|}$ by

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

2

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[1]. 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 DiPerna and Lions in [72].

Let us now make some basic technical remarks on the continuity equation and the transport equation:

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

$$\int_I \int_A |b_t| \, d|\mu_t| \, dt < +\infty \qquad \forall A \Subset \mathbb{R}^d \tag{1}$$

(here and in the sequel, $A \Subset B$ means that $\overline{A}$ is a compact subset of B). However, when we consider possibly singular measures μ_t, we must take care of the fact that the product $b_t \mu_t$ is sensitive to modifications of 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 $b_t \cdot \nabla w$ is defined by

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

(a definition consistent with the case when w_t is smooth) the transport equation makes sense *only* if we assume that $D_x \cdot b_t = \operatorname{div} b_t \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in I$. See also [41, 44] for more refined results on the transport equation when b satisfies a one-sided Lipschitz condition.

---

[1] In these lectures we are using the notation $f_\# \mu$ for the push-forward operator between measures, see (3) and (4).

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, if $\mu_t \geq 0$ 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 $c_t = -w_t \text{div} \, \boldsymbol{b}_t$.
(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 [41, 44]).
(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 [72] to characterize the flow $\boldsymbol{Y}$ and to prove its stability properties. The approach developed here, based on [10], is based on a careful analysis of the measures transported by the flow, and ultimately on the homogeneous continuity equation only.

Plan of these lecture notes. In Section 2 we review the Cauchy-Lipschitz theory and we derive classical representation formulas for the solution of the continuity and the transport

equations in the Lipschitz regularity setting. Section 3 is devoted to general principles, somewhat extending the classical theory of characteristics to the non-smooth setting, relating the uniqueness of the ODE to the uniqueness of the PDE. The main tool here is the superposition principle, asserting that every positive measure solution to the continuity equation can indeed be realised as a suitable "probabilistic push-forward" of the initial datum along a (possibly multi-valued) ODE flow. We then introduce the notion of Lagrangian flow (the suitable notion of solution to the ODE in the non-smooth framework) and prove that the well posedness of the PDE with velocity field $\boldsymbol{b}$ implies the well posedness of the Lagrangian flow associated to b.

We then start the study of the PDE in various weak regularity contexts: in Section 4 we introduce the notion of renormalized solution, prove the well posedness of the PDE for Sobolev vector fields, and describe the consequences for the Lagrangian flow; this is extended to vector fields with bounded variation in Section 5.

In the subsequent two sections we describe an alternative approach to the theory of Lagrangian flows based on suitable quantitative estimates, without reference to the PDE theory: this is first carried out for Sobolev $W^{1,p}$ vector fields with $p > 1$ in Section 6, by showing how suitable bounds on an integral quantity measuring a logarithmic distance between Lagrangian flows imply well-posedness, quantified stability and mild regularity for the Lagrangian flow, and then extended in Section 7 to $W^{1,1}$ vector fields, and to vector fields whose derivative is given by a singular integral of an L^1 function.

In Section 8 we summarise two applications to nonlinear PDEs (the Keyfitz-Kranzer system of conservation laws and the semi-geostrophic equations), and in the final Section 9 we collect some open problems and comment on further literature on the subject.

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. Quantitative information can be obtained by differentiation:

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

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

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

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

Given metric spaces X, Y and a Borel map $f : X \to Y$, we shall use in these lectures the notation $f_\#$ for the push-forward operator mapping nonnegative Borel measures in X to nonnegative Borel measures in Y, namely

$$f_\# \mu(B) := \mu\big(f^{-1}(B)\big) \qquad B \subset Y \text{ Borel.} \tag{3}$$

Notice that $f_\#$ is mass-preserving, so that it maps finite Borel measures to finite Borel measures and probability measures to probbility measures. Notice also the basic change of variables formula

$$\int_Y \phi \, df_\# \mu = \int_X \phi \circ f \, d\mu \tag{4}$$

for $\phi : Y \to [0, +\infty]$ Borel, that we will frequently and tacitly use in the sequel.

The representation formula for solutions to the continuity equation is indeed very simple:

Proposition 4 *For any probability measure $\bar{\mu}$ in $\mathbb{R}^d$ the solution of the continuity equation is given by*

$$\mu_t := \boldsymbol{X}(t, \cdot)_\# \bar{\mu}, \quad i.e. \quad \int_{\mathbb{R}^d} \varphi \, d\mu_t = \int_{\mathbb{R}^d} \varphi(\boldsymbol{X}(t, x)) \, d\bar{\mu}(x). \tag{5}$$

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

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

$$\leq \|\nabla\varphi\|_\infty \int_{\cup_i(a_i,b_i)} \sup |\boldsymbol{b}_t| \, dt$$

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

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

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 dominated convergence gives

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

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

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 *explicitly* 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:

$$
f_\#(\rho \mathscr{L}^d) = \frac{\rho}{|Jf|} \circ f^{-1} \mathscr{L}^d. \tag{6}
$$

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

$$
\exp\left[-\int_0^t \|[\operatorname{div} \boldsymbol{b}_s]^-\|_\infty\, ds \right] \leq J\boldsymbol{X}(t,x) \leq \exp\left[\int_0^t \|[\operatorname{div} \boldsymbol{b}_s]^+\|_\infty\, ds \right] \tag{7}
$$

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 linearized ODE $\frac{d}{dt}\nabla\boldsymbol{X} = \nabla\boldsymbol{b}_t(\boldsymbol{X})\nabla\boldsymbol{X}$.

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

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

with respect to the w^*-topology of $W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d)$ (see for instance Section 8.2.4 of [75]). 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*

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

in the sense of distributions, 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 $L^1(W_{\text{loc}}^{1,\infty})$ assumption, and assuming $\mu_t \geq 0$, the second growth condition on $|\boldsymbol{b}|$ can be replaced by a global, but more intrinsic, condition:

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_t|}{1+|x|}\, d\mu_t\, dt < +\infty. \tag{8}$$

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 $t \in [0, T]$, see Section 8.1 of [16].

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 [106] of generalized surfaces and controls (a theory with remarkable applications also to non-linear PDE's [71, 99] and Calculus of Variations [27]) and has also some connection with Brenier's weak solutions of incompressible Euler equations [37, 21], with Kantorovich's viewpoint in the theory of optimal transportation [76, 95] and with Mather's theory [90, 91, 26]. 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 in the class*

$$\left\{ \mu_t : \int_0^T \int \frac{|\boldsymbol{b}_t|}{1+|x|}\, d\mu_t dt < \infty \right\}$$

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 $\delta_{\tilde{\boldsymbol{X}}(t)}$.

9

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. $\qquad\square$

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

$$\int_0^T \int_{\mathbb{R}^d \times \Gamma_T} \chi_{B_R}(e_t)|\boldsymbol{b}_t(e_t)| \, d\boldsymbol{\eta} \, dt < +\infty \qquad \forall R > 0 \tag{9}$$

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

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

for $\boldsymbol{\eta}$-a.e. (x, γ), with R such that $\mathrm{supp}\,\varphi \subset \overline{B}_R$. Therefore, assuming for simplicity that $\boldsymbol{\eta}(\mathbb{R}^d \times \Gamma_T) = 1$, an integration with respect to $\boldsymbol{\eta}$ gives

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

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

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

10

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

$$\int_{\mathbb{R}^d \times \Gamma_T} \varphi(x, \gamma) \, d\boldsymbol{\eta}(x, \gamma) := \int_{\mathbb{R}^d} \left(\int_{\Gamma_T} \varphi(x, \gamma) \, d\boldsymbol{\eta}_x(\gamma) \right) d\bar{\mu}(x). \tag{10}$$

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*

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

Then μ_t is a superposition solution, i.e. there exists $\boldsymbol{\eta} \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ concentrated on pairs (x, γ) solution to the ODE such that $\mu_t = \mu_t^{\boldsymbol{\eta}}$ for any $t \in [0, T]$. In particular $t \mapsto \mu_t(\mathbb{R}^d)$ is constant in $[0, T]$.

In the proof we use the *weak* 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 weak convergence. Remember that tightness means:

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

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

Proof. **Step 1 (smoothing).** [77] 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

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

11

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

$$\int_{\mathbb{R}^d} \varphi \, d\mu_t^{\boldsymbol{\eta}_\epsilon} = \int_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t)) \, d\boldsymbol{\eta}^\epsilon \tag{11}$$

$$= \int_{\mathbb{R}^d} \varphi(\boldsymbol{X}^\epsilon(t, x)) \, d\mu_0^\epsilon(x) = \int_{\mathbb{R}^d} \varphi \, d\mu_t^\epsilon.$$

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

$$((1+|x|)c) * \rho_\epsilon \le (1+|x|)c * \rho_\epsilon + \epsilon c * \tilde{\rho}_\epsilon \tag{12}$$

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

$$\Theta(|\boldsymbol{b}_t^\epsilon(x)|)\mu_t^\epsilon(x) \le (\Theta(|\boldsymbol{b}_t|)\mu_t) * \rho_\epsilon(x). \tag{13}$$

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 $z \ne 0$, and to 0 if $z = t = 0$) and with the measure $\rho_\epsilon(x - \cdot)\mathscr{L}^d$.

Let us introduce the functional

$$\Psi(x, \gamma) := \begin{cases} \psi(x) + \int_0^T \frac{\Theta(|\dot\gamma|)}{1+|\gamma|} \, dt & \text{if } \gamma \in AC([0,T]; \mathbb{R}^d) \text{ and } x = \gamma(0); \\ +\infty & \text{otherwise.} \end{cases}$$

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 $\{\Psi \le t\}$, so that also $\int_0^T \Theta(|\dot\gamma|) \, dt$ is uniformly bounded on the sublevels of Ψ). Since

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

$$\overset{(12),(13)}{\le} (1 + \epsilon M) \int_0^T \int_{\mathbb{R}^d} \frac{\Theta(|\boldsymbol{b}_t|(x))}{1+|x|} \, d\mu_t dt$$

and

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

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 weakly sequentially relatively compact as $\epsilon \downarrow 0$. If $\boldsymbol{\eta}$ is any limit point we can pass to the limit in (11) 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

$$\int_{\mathbb{R}^d \times \Gamma_T} \frac{\left| \gamma(t) - x - \int_0^t \boldsymbol{b}_s(\gamma(s))\, ds \right|}{1 + \max_{[0,T]} |\gamma|} \, d\boldsymbol{\eta} = 0 \tag{14}$$

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

$$\int_{\mathbb{R}^d \times \Gamma_T} \frac{\left| \gamma(t) - x - \int_0^t \boldsymbol{c}_s(\gamma(s))\, ds \right|}{1 + \max_{[0,T]} |\gamma|} \, d\boldsymbol{\eta} \leq \int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_s - \boldsymbol{c}_s|}{1 + |x|} \, d\mu_s ds \tag{15}$$

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

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

and noticing that

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

we can pass to the limit in (15) with $\boldsymbol{c} = \boldsymbol{c}^n$ to obtain (14).

It remains to show (15). This is a limiting argument based on the fact that (14) holds for $\boldsymbol{b}^\epsilon$, $\boldsymbol{\eta}^\epsilon$:

$$\int_{\mathbb{R}^d \times \Gamma_T} \frac{\left| \gamma(t) - x - \int_0^t \boldsymbol{c}_s(\gamma(s))\, ds \right|}{1 + \max_{[0,T]} |\gamma|} \, d\boldsymbol{\eta}^\epsilon$$

$$= \int_{\mathbb{R}^d} \frac{\left| \boldsymbol{X}^\epsilon(t, x) - x - \int_0^t \boldsymbol{c}_s(\boldsymbol{X}^\epsilon(s, x))\, ds \right|}{1 + \max_{[0,T]} |\boldsymbol{X}^\epsilon(\cdot, x)|} \, d\mu_0^\epsilon(x)$$

$$= \int_{\mathbb{R}^d} \frac{\left| \int_0^t \boldsymbol{b}_s^\epsilon(\boldsymbol{X}^\epsilon(s, x)) - \boldsymbol{c}_s(\boldsymbol{X}^\epsilon(s, x))\, ds \right|}{1 + \max_{[0,T]} |\boldsymbol{X}^\epsilon(\cdot, x)|} \, d\mu_0^\epsilon(x) \leq \int_0^t \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_s^\epsilon - \boldsymbol{c}_s|}{1 + |x|} \, d\mu_s^\epsilon ds$$

$$\leq \int_0^t \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_s^\epsilon - \boldsymbol{c}_s^\epsilon|}{1 + |x|} \, d\mu_s^\epsilon ds + \int_0^t \int_{\mathbb{R}^d} \frac{|\boldsymbol{c}_s^\epsilon - \boldsymbol{c}_s|}{1 + |x|} \, d\mu_s^\epsilon ds$$

$$\leq \int_0^t \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_s - \boldsymbol{c}_s|}{1 + |x|} \, d\mu_s ds + \int_0^t \int_{\mathbb{R}^d} \frac{|\boldsymbol{c}_s^\epsilon - \boldsymbol{c}_s|}{1 + |x|} \, d\mu_s^\epsilon ds.$$

13

In the last inequalities we added and subtracted $\boldsymbol{c}_t^\epsilon := (\boldsymbol{c}_t \mu_t) * \rho_\epsilon / \mu_t^\epsilon$. Since $\boldsymbol{c}_t^\epsilon \to \boldsymbol{c}_t$ uniformly as $\epsilon \downarrow 0$ thanks to the uniform continuity of $\boldsymbol{c}$, passing to the limit in the chain of inequalities above we obtain (15). $\qquad\square$

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

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

$$0 \leq \mu_t' \leq \mu_t \in \mathscr{L}_{\boldsymbol{b}} \qquad \Longrightarrow \qquad \mu_t' \in \mathscr{L}_{\boldsymbol{b}} \tag{16}$$

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 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 well-posedness for the PDE holds in the class $\mathscr{L}_{\boldsymbol{b}}$.

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

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

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

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

$$\boldsymbol{\eta}_x\left(\{\gamma: \; \gamma(t) \in E\}\right)\boldsymbol{\eta}_x\left(\{\gamma: \; \gamma(t) \in E'\}\right) = 0 \quad \bar{\mu}\text{-a.e. in } \mathbb{R}^d. \tag{17}$$

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 (17) 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 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 it all other sets $\{\lambda_x(R_j) > 0\}$, $j \neq i$). Since δ is arbitrary the statement is proved. $\square$

Theorem 16 (Uniqueness of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows) *Assume that the PDE is well posed 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. But this is not possible, since it would contradict Theorem 18. Such measure η can be easily 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 $\mu_t^{\boldsymbol{\eta}} = \frac{1}{2}(\mu_t^{\boldsymbol{\eta}^1} + \mu_t^{\boldsymbol{\eta}^2}) \in \mathscr{L}_{\boldsymbol{b}}$. $\square$

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

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

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

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

Theorem 18 *Assume that the PDE is well posed 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*

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

15

Proof. We use the representation (10) of $\boldsymbol{\eta}$, given by the disintegration theorem, the criterion stated in Lemma 15 and argue by contradiction. If the thesis were false then $\boldsymbol{\eta}_x$ would not be a Dirac mass in a set of $\bar{\mu}$-positive measure and we could 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

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

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

$$0 < \frac{\boldsymbol{\eta}_x(\{\gamma : \ \gamma(t) \in E\})}{\boldsymbol{\eta}_x(\{\gamma : \ \gamma(t) \in E'\})} \le M \qquad \forall x \in C \tag{18}$$

for some constant $M < +\infty$. Denoting by β the quotient in (18), we define measures $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2$ whose disintegrations $\boldsymbol{\eta}_x^1$, $\boldsymbol{\eta}_x^2$ are given by

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

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

$$\mu_0^1 = \boldsymbol{\eta}_x(\{\gamma : \ \gamma(t) \in E\})\bar{\mu}\llcorner C, \qquad \mu_0^2 = \beta(x)\boldsymbol{\eta}_x(\{\gamma : \ \gamma(t) \in E'\})\bar{\mu}\llcorner C,$$

so that the definition of β yields $\mu_0^1 = \mu_0^2$. On the other hand, μ_t^1 is orthogonal to μ_t^2: precisely, denoting by $\boldsymbol{\eta}_{tx}$ the image of $\boldsymbol{\eta}_x$ under the map $\gamma \mapsto \gamma(t)$, we have

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

Notice also that $\mu_t^1 \le \mu_t$, $\mu_t^2 \le \mu_t$ and so the monotonicity assumption (16) on $\mathscr{L}_{\boldsymbol{b}}$ gives $\mu_t^i \in \mathscr{L}_{\boldsymbol{b}}$. This contradicts the assumption on the validity of the comparison principle in $\mathscr{L}_{\boldsymbol{b}}$. $\quad\square$

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 is well posed 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

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

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. $\quad\square$

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

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

whenever $\mu_t^h \in \mathscr{L}_{\boldsymbol{b}^h}$ and $\mu_t^h \to \mu_t$ weakly 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_{loc}.

For the flows, the natural convergence in 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 $\bar{\mu}$-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*

$$\mu_t^h \to \mu_t \qquad \text{weakly as } h \to \infty \text{ for all } t \in [0,T] \tag{19}$$

$$\limsup_{h \to \infty} \int_0^T \int_{\mathbb{R}^d} \frac{\Theta(|\boldsymbol{b}_t^h|)}{1+|x|} \, d\mu_t^h dt \le \int_0^T \int_{\mathbb{R}^d} \frac{\Theta(|\boldsymbol{b}_t|)}{1+|x|} \, d\mu_t dt < +\infty \tag{20}$$

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

(iv) *the PDE is well posed 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.

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

and we obtain, using (20) 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 $\boldsymbol{\eta}^h$ in the weak topology. Using the same argument used in Step 3 of the proof of the superposition principle and (20) 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 $\mu_t^h \to \mu_t$ weakly 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

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

By applying the following general principle we conclude. $\qquad \square$

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

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

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 $\bar{\mu}(\{v \neq w\}) \leq \epsilon$. We define

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

and notice that

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

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

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

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

and since ϵ is arbitrary the proof is achieved. $\qquad \square$

Lemma 23 *Let $A \subset \mathbb{R}^m$ be an open set, and let $\sigma^h \in \mathscr{M}_+(A)$ be weakly 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) $\displaystyle\limsup_{h \to \infty} \int_A \Theta(|\boldsymbol{f}^h|) \, d\sigma^h \leq \int_A \Theta(|\boldsymbol{f}|) \, d\sigma < +\infty$ 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 weak 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

$$\int_{A \times \mathbb{R}^k} \psi \, d\nu \leq \liminf_{h \to \infty} \int_{A \times \mathbb{R}^k} \psi \, d\nu_h. \tag{21}$$

Then, choosing test functions $\psi = \psi(x) \in C_b(A)$, the weak 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

$$\int_{A \times \mathbb{R}^k} \psi(x, y) \, d\nu(x, y) = \int_A \left(\int_{\mathbb{R}^k} \psi(x, y) \, d\nu_x(y) \right) d\sigma(x) \tag{22}$$

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

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

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 (21), assumption (ii) gives

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

hence $\int \Theta(|y|) \, d\nu_x = \boldsymbol{f}(x) = \Theta(|\int y \, d\nu_x|)$ 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 (22) of ν with $\nu_x = \delta_{\boldsymbol{f}(x)}$, the convergence statement can be achieved just choosing the test function $\psi(x, y) = |y - \boldsymbol{c}(x)|$. $\qquad\square$

4 Vector fields with Sobolev spatial regularity

Here we discuss the well-posedness of the continuity or transport equations assuming the $\boldsymbol{b}_t(\cdot)$ has Sobolev regularity, following [72]. 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 $\boldsymbol{b} \in L^1_{\text{loc}}\left(I; L^1_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^d)\right)$ 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*

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

Let $w \in L^\infty_{\text{loc}}\left(I; L^\infty_{\text{loc}}(\mathbb{R}^d)\right)$ and assume that, in the sense of distributions, there holds

$$c := \frac{d}{dt}w + \boldsymbol{b} \cdot \nabla w \in L^1_{\text{loc}}(I \times \mathbb{R}^d). \tag{23}$$

Then, we say that w is a renormalized solution of (23) *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

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

Notice also that the concept makes sense, choosing properly the class of "test" functions β, also for functions w that do not satisfy (23), not even locally integrable. This is particularly relevant in connection with DiPerna-Lions's existence theorem for Boltzmann's equation [73], 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:

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

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 $w^* - L^\infty_{\text{loc}}$ 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_{loc} topology at almost every time. 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 [103, 92]. We remark the fact that, in general, a renormalized solution does *not* need to have a representative which is strongly continuous *for every* t. This can be seen using a variation of an example given by Depauw [68]. Depauw's example provides a divergence free vector field $\boldsymbol{a} \in L^\infty([0,1] \times \mathbb{R}^2; \mathbb{R}^2)$, with $\boldsymbol{a}(t, \cdot) \in BV_{\text{loc}}(\mathbb{R}^2; \mathbb{R}^2)$ for $\mathscr{L}^1$-a.e. $t \in [0,1]$ (but $\boldsymbol{a} \notin L^1([0,1]; BV_{\text{loc}})$) such that the Cauchy problem

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

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

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

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

Remark 26 A new insight in the theory of renormalized solutions has been obtained in [34]. In particular, it is proved that for a vector field $\boldsymbol{b} \in L^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ with zero divergence (and without any regularity assumption) the following two conditions are equivalent (the L^2 framework has been considered just for simplicity):

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

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

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

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

The proof of this equivalence is obtained through the study of the approximation properties of the solution of the transport equation, with respect to the norm of the graph of the transport operator (see Theorem 2.1 of [34] for the details).

Using the concept of renormalized solution we can prove a comparison principle (and therefore well-posedness) in the following natural class $\mathscr{L}$:

$$\mathscr{L} := \left\{ w \in L^\infty\left([0,T]; L^1(\mathbb{R}^d)\right) \cap L^\infty\left([0,T]; L^\infty(\mathbb{R}^d)\right) : \right. \tag{24}$$

$$\left. w \in C\left([0,T]; w^* - L^\infty(\mathbb{R}^d)\right) \right\}.$$

Theorem 27 (Comparison principle) *Assume that*

$$\frac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right) + L^1\left([0,T]; L^1(\mathbb{R}^d)\right), \tag{25}$$

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

$$[\operatorname{div}\boldsymbol{b}_t]^- \in L^1_{\mathrm{loc}}\left([0,T) \times \mathbb{R}^d\right). \tag{26}$$

Setting $\boldsymbol{b}_t \equiv 0$ for $t < 0$, assume in addition that any *solution of* (23) *in $(-\infty, T) \times \mathbb{R}^d$ is renormalized. Then the comparison principle, and therefore uniqueness, for the continuity equation hold in the class $\mathscr{L}$ defined in* (24).

Proof. By the linearity of the equation, it suffices to show that $w \in \mathscr{L}$ and $w_0 \leq 0$ implies $w_t \leq 0$ 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 $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi \subset \overline{B}_2(0)$ and $\varphi \equiv 1$ on $B_1(0)$, and the renormalization functions

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

Notice that

$$\beta_\epsilon(t) \uparrow t^+ \quad \text{as } \epsilon \downarrow 0, \qquad t\beta_\epsilon'(t) - \beta_\epsilon(t) \in [0, \epsilon]. \tag{27}$$

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 $R \geq 1$, into the PDE we obtain

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

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

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

and using the inequality

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

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

$$3\|\nabla\varphi\|_\infty\|\frac{\boldsymbol{b}_{1t}}{1+|x|}\|_\infty \int_{\{|x|\geq R\}} |w_t|\, dx + 3\|\nabla\varphi\|_\infty\|w_t\|_\infty \int_{\{|x|\geq R\}} \frac{|\boldsymbol{b}_{2t}|}{1+|x|}\, dx.$$

The second integral can be estimated with

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

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

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

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

Remark 28 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, DiPerna and Lions proved that all distributional solutions are renormalized when there is Sobolev regularity with respect to the spatial variables.

Theorem 29 *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 (23). 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). It will be proven in the next proposition that this is indeed the case, and it is exactly here that the Sobolev regularity plays a role. $\qquad\square$

Proposition 30 (Strong convergence of commutators) *If* $w \in L^\infty_{\text{loc}}(I \times \mathbb{R}^d)$ *and* $\boldsymbol{b} \in L^1_{\text{loc}}\left(I; W^{1,1}_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^d)\right)$ *we have*

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

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

$$r^\epsilon(t, x) = \int_{\mathbb{R}^d} w(t, y)(\boldsymbol{b}_t(y) - \boldsymbol{b}_t(x)) \cdot \nabla \rho_\epsilon(x - y)\, dy - (w \text{div}\, \boldsymbol{b}_t) * \rho_\epsilon(x). \tag{28}$$

Indeed, to prove (28) we introduce 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 $w^\epsilon := w * \rho_\epsilon$). In order to get (28) it suffices to show that $R^\epsilon = L^\epsilon - w^\epsilon \text{div}\, \boldsymbol{b}_t$, where

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

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

$$
\begin{aligned}
& -\int_I \int w\boldsymbol{b} \cdot \nabla \rho_\epsilon * \varphi dydt - \int_I \int \varphi \boldsymbol{b} \cdot \nabla \rho_\epsilon * w dx dt - \int_I \int w^\epsilon \varphi \text{div}\, \boldsymbol{b}_t dt \\
= \quad & -\int_I \int \int w_t(y)\boldsymbol{b}_t(y) \cdot \nabla \rho_\epsilon(y - x)\varphi(x)\, dx dy dt \\
& -\int_I \int \int \boldsymbol{b}_t(x)\nabla \rho_\epsilon(x - y)w_t(y)\varphi(x) dy dx dt - \int_I \int w^\epsilon \varphi \text{div}\, \boldsymbol{b}_t\, dx dt \\
= \quad & \int_I \int L^\epsilon \varphi\, dx dt - \int_I \int w^\epsilon \varphi \text{div}\, \boldsymbol{b}_t\, dx dt
\end{aligned}
$$

(in the last equality we used the fact that $\nabla \rho$ is odd).
Now we change variables and write (28) in the form

$$r^\epsilon(t, x) = \int_{\mathbb{R}^d} w(t, x - \epsilon z)\frac{(\boldsymbol{b}_t(x - \epsilon z) - \boldsymbol{b}_t(x)) \cdot \nabla \rho(z)}{\epsilon}\, dz - (w \text{div}\, \boldsymbol{b}_t) * \rho_\epsilon(x). \tag{29}$$

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)

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

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

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

The elementary identity

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

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, so the main point here is the strength of the convergence).

□

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

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

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

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

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

The assumption (b) expresses in a quantitative way the condition that "trajectories do not concentrate too much". It can be equivalently stated as

$$\int_{\mathbb{R}^d} \varphi(\boldsymbol{X}(t, x))\, dx \le C \int_{\mathbb{R}^d} \varphi(y)\, dy$$

for all positive $\varphi \in C_c(\mathbb{R}^d)$.

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

Theorem 31 (Existence and uniqueness of $\mathscr{L}$-Lagrangian flows) *Let* $\boldsymbol{b} \in L^1\left([0, T]; W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^d)\right)$ *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) $[\mathrm{div}\, \boldsymbol{b}_t]^- \in L^1\left([0, T]; L^\infty(\mathbb{R}^d)\right).$

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

Proof. By the previous results, well-posedness 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

$$\frac{d}{dt}w + D_x \cdot (\boldsymbol{b}w) = 0 \tag{30}$$

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_{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 (30). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

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 [72]. Heuristically, while the DiPerna-Lions one is based on the semigroup of transformations $x \mapsto \boldsymbol{X}(t, x)$, ours is based on the properties of the map $x \mapsto \boldsymbol{X}(\cdot, x)$.

Remark 32 The definition of the flow in [72] 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

$$\tilde{\boldsymbol{X}}(t, x) := \begin{cases} \boldsymbol{X}^{s'}(t, x) & \text{if } t \in [s', s]; \\[2mm] \boldsymbol{X}^s\left(t, \boldsymbol{X}^{s'}(s, x)\right) & \text{if } t \in [s, T]. \end{cases}$$

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 31 (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 [13] for more general results).

Theorem 33 (Stability) *Let $\boldsymbol{b}^h$, $\boldsymbol{b} \in L^1\left([0,T]; W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^d)\right)$, 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 $L^1_{\mathrm{loc}}\left((0,T) \times \mathbb{R}^d\right)$;

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

(iii) $[\mathrm{div}\,\boldsymbol{b}^h_t]^-$ 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 compact support. Under this assumption, (i) and (iii) ensure that $\mu^h_t \leq M \chi_{B_R} \mathscr{L}^d$ for some constants M and R independent of h and t. Denoting by μ_t the weak limit of μ^h_t, choosing $\Theta(z) = |z|^2$ in (iii) of Theorem 21, we have to check that

$$\lim_{h\to\infty} \int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_h|^2}{1+|x|}\, d\mu^h_t dt = \int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}|^2}{1+|x|}\, d\mu_t dt. \tag{31}$$

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 $|\tilde{\boldsymbol{b}}_\epsilon| \leq C$ and $\tilde{\boldsymbol{b}}_\epsilon = \boldsymbol{b}$ on $[0,T] \times B_R \setminus B$. We write

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

so that

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

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 (31) is fulfilled. $\qquad\square$

Finally, we conclude this section with the illustration of some recent results [83, 17, 18] that seem to be more specific of the Sobolev case, concerned with the "differentiability" w.r.t. to x of the flow $\boldsymbol{X}(t, x)$, see [13] for a more detailed treatment of this topic. These results provide a sort of bridge with the standard Cauchy-Lipschitz calculus:

Theorem 34 *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

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

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. Also compare with the approximate differentiability results presented in Section 6.5.

5 Vector fields with BV spatial regularity

In this section we prove the renormalization Theorem 29 under the weaker assumption of a BV dependence w.r.t. the spatial variables, but still assuming that

$$D \cdot \boldsymbol{b}_t \ll \mathscr{L}^d \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0, T). \tag{32}$$

Theorem 35 *Let $\boldsymbol{b} \in L^1_{\mathrm{loc}}\left((0, T); BV_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^d)\right)$ be satisfying (32). Then any distributional solution $w \in L^\infty_{\mathrm{loc}}\left((0, T) \times \mathbb{R}^d\right)$ of*

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

is a renormalized solution.

We try to give a reasonably detailed proof of this result, referring to the original paper [10] 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 space-time 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 homogeneity and locality of the arguments involved, that $\|w\|_\infty \leq 1$. 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

$$r^\epsilon(t, x) = \int_{\mathbb{R}^d} w(t, x - \epsilon y) \frac{(\boldsymbol{b}_t(x - \epsilon y) - \boldsymbol{b}_t(x)) \cdot \nabla \rho(y)}{\epsilon} \, dy - (w\mathrm{div}\, \boldsymbol{b}_t) * \rho_\epsilon(x) \tag{33}$$

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 $u \in BV_{\mathrm{loc}}$ and any $z \in \mathbb{R}^d$ with $|z| < \epsilon$ we have a classical L^1 estimate on the difference quotients

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

where $Du = (D_1 u, \dots, D_d u)$ stands for the distributional derivative of u, $D_z u = \langle Du, z \rangle = \sum_i z_i D_i u$ denotes the component along z of Du and K_ϵ is the open ϵ-neighbourhood of K. Its proof follows from an elementary smoothing and lower semicontinuity argument.

We notice that, setting $D\boldsymbol{b}_t = M_t |D\boldsymbol{b}_t|$, we have

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

and therefore (using that $\mathrm{div}\, \boldsymbol{b}_t \mathscr{L}^d$ is the trace of $D^a \boldsymbol{b}_t$) the L^1 estimate on difference quotients gives the *anisotropic estimate*

$$\limsup_{\epsilon \downarrow 0} \int_K |r^\epsilon| \, dx \leq \int_K \int_{\mathbb{R}^d} |\langle M_t(x) z, \nabla \rho(z) \rangle| \, dz \, d|Db|(t,x) + d|D^a b|(K) \qquad (34)$$

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 $b(t, \cdot) \in W^{1,1}_{\mathrm{loc}}$ 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

$$\int_K |\hat{\mu}_\epsilon| \, dt \leq |\mu|(K_\epsilon) \quad \text{for any compact set } K \subset \mathbb{R}, \qquad (35)$$

where K_ϵ is again the open ϵ neighbourhood of K. A density argument based on (35) then shows that $\hat{\mu}_\epsilon$ converge in $L^1_{\mathrm{loc}}(\mathbb{R})$ to the density of μ with respect to $\mathscr{L}^1$ whenever $\mu \ll \mathscr{L}^1$. If $u \in BV_{\mathrm{loc}}$ 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_{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 $L^1_{\text{loc}}(\mathbb{R}^d)$ 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*

$$\limsup_{\epsilon\downarrow 0} \int_K |r^\epsilon|\, dx \le \left(\int_K \int_{\mathbb{R}^d} |z||\nabla\rho(z)|\, dz \right) d|D^s\boldsymbol{b}|(t,x) \tag{36}$$

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 $|D^sb|(K) \ll |D^ab|(K)$, while the anisotropic one turns out to be useful in the regions where the dominant part is the singular one, i.e. $|D^ab|(K) \ll |D^sb|(K)$. 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

$$\frac{d}{dt}\beta(w^\epsilon) + \boldsymbol{b} \cdot \nabla\beta(w^\epsilon) - \beta'(w^\epsilon)c * \rho_\epsilon = \beta'(w^\epsilon)r^\epsilon. \tag{37}$$

Let L be the supremum of $|\beta'|$ on $[-1,1]$. Then, since K is an arbitrary compact set, (36) 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 (34) tells also us that

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

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

$$\nu \le L \int_{\mathbb{R}^d} |\langle M.(\cdot)z, \nabla\rho(z)\rangle|\, dz|D^s\boldsymbol{b}|. \tag{38}$$

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 (37) 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

$$|\sigma| \le L\Lambda(M.(\cdot), \rho)|D^s\boldsymbol{b}| \quad \text{with} \quad \Lambda(N,\rho) := \int_{\mathbb{R}^d} |\langle Nz, \nabla\rho(z)\rangle|\, dz. \tag{39}$$

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

$$G(N) := \inf \left\{ \Lambda(N, \rho) : \ \rho \in C_c^\infty(B_1), \ \rho \geq 0, \ \int_{\mathbb{R}^d} \rho = 1 \right\} \tag{40}$$

with $N = M_t(x)$. Indeed, notice that (39) gives

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

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

$$|\sigma| \leq LG(M_\cdot(\cdot))|D^s \boldsymbol{b}|. \tag{41}$$

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

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

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 (41) we obtain that $\sigma = 0$, thus concluding the proof.

Lemma 36 (Alberti) *For any $d \times d$ matrix N the infimum in* (40) *is* $|\text{trace } N|$.

Proof. Heuristically, in the case when N is traceless, 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

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

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

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

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

$$\int_{\mathbb{R}^d} |\text{div } (Nz\rho(z))| \, dz \leq \frac{2}{T}. \tag{42}$$

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

$$\rho(z) := \frac{1}{T} \int_0^T \theta(e^{-tN}z)e^{-t \text{ trace } N} \, dt \tag{43}$$

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

31

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

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

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

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

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

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

But this is elementary, since

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

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

The original argument in [10] was slightly different and used, instead of Lemma 36, a much deeper result, still due to Alberti, saying that for a BV_{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 31 the assumption $\boldsymbol{b} \in L^1\left([0, T]; W_{\mathrm{loc}}^{1,1}(\mathbb{R}^d; \mathbb{R}^d)\right)$ with $\boldsymbol{b} \in L^1\left([0, T]; BV_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^d)\right)$, 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 33 (for $\boldsymbol{b}$ and $\boldsymbol{b}^h$) we obtain *stability* of $\mathscr{L}$-Lagrangian flows.

6 Quantitative ODE estimates: $W^{1,p}$ regularity with $p > 1$

In this and in the next section we introduce a further approach to the well-posedness of the ODE, based on quantitative a priori estimates, as an alternative to the approach exploiting

the well-posedness of the continuity equation (in the various regularity contexts) and the link between the ODE and the continuity equation, described in Section 3.

We restrict here our attentions to $\mathscr{L}$-Lagrangian flows with $\mathscr{L} := L^\infty(L^\infty)$, and for simplicity we refer to them (in this and in the next section) simply as regular Lagrangian flows. We recall that, given a vector field $\boldsymbol{b} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, we are considering as "admissible solutions to the ODE" those maps $\boldsymbol{X} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that:

(i) For $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$, the function $t \mapsto \boldsymbol{X}(t, x)$ is a solution of the ODE in the integral sense;

(ii) There exists a constant $L > 0$ such that $X(t, \cdot)_\# \mathscr{L}^d \leq L \mathscr{L}^d$ for all $t \in [0, T]$.

The constant L in (ii) is called *compressibility constant* of the flow $\boldsymbol{X}$.

For simplicity, in this and in the next section we only consider globally bounded vector fields: the extension to more general growth conditions does not pose any problem, but makes some of the computations much longer. At many occurrences we will also avoid to explicitly indicate the time dependence. We deal with three different regularity classes: $W^{1,p}$ with $p > 1$ (in this section), $W^{1,1}$ and the case in which the derivative is a singular integral of an L^1 function (in the next section). We remark that we do *not* systematically assume conditions about the spatial divergence div $\boldsymbol{b}$: divergence bounds will be needed for the existence result only.

The objective is to develop in the non-Lipschitz context a theory for the ODE which is "parallel" to the usual Cauchy-Lipschitz theory for the Lipschitz context, and which is independent from the PDE theory. In particular, this approach allows to recover for the ODE:

- Existence and uniqueness;

- Stability (with respect to approximations of the vector field), with an explicit convergence rate;

- Compactness, under natural bounds (this is relevant in view of applications to nonlinear PDEs);

- Regularity of the flow $\boldsymbol{X}(t, x)$ with respect to the initial position x (this also gives some mild propagation of regularity for the transport equation).

The presentation is based on [61, 36]. Some ideas originated from [17, 83].

6.1   An integral quantity and the lower bound

Remember the very basic estimate providing (2), in the context of Lipschitz vector fields. A variant of it can be performed using the following formal estimate for the logarithm of the spatial derivative of the flow:

$$\frac{d}{dt} \log |\nabla \boldsymbol{X}| \leq \frac{1}{|\nabla \boldsymbol{X}|} \left| \frac{d}{dt} \nabla \boldsymbol{X} \right| = \frac{1}{|\nabla \boldsymbol{X}|} |\nabla (\boldsymbol{b}(\boldsymbol{X}))| = |\nabla \boldsymbol{b}|(\boldsymbol{X}). \tag{44}$$

Again, when $\boldsymbol{b}$ is Lipschitz, we recover that the flow is Lipschitz, with a Lipschitz constant depending exponentially on the Lipschitz constant of $\boldsymbol{b}$.

The relevance of (44) is that it admits an integral version, which is working well even out of the Lipschitz case, in some weak regularity contexts.

Let us fix two regular Lagrangian flows $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, associated to the vector fields $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, and with compressibility constants L_1 and L_2. Given a (small) parameter $\delta > 0$ and a truncation radius $R > 0$, we consider the following (time dependent) quantity:

$$\Phi_\delta(t) = \int_{B_R} \log\left(1 + \frac{|\boldsymbol{X}_1(t,x) - \boldsymbol{X}_2(t,x)|}{\delta}\right) dx. \tag{45}$$

We start by establishing a simple but useful lower bound for Φ_δ. Given an arbitrary $\gamma > 0$ we can compute

$$\Phi_\delta(t) \geq \int_{B_R \cap \{|\boldsymbol{X}_1 - \boldsymbol{X}_2| > \gamma\}} \log\left(1 + \frac{\gamma}{\delta}\right) dx = \mathscr{L}^d\left(B_R \cap \{|\boldsymbol{X}_1 - \boldsymbol{X}_2| > \gamma\}\right) \log\left(1 + \frac{\gamma}{\delta}\right). \tag{46}$$

In particular, in the special case $\boldsymbol{b}_1 = \boldsymbol{b}_2 = \boldsymbol{b}$ (in which Φ_δ is somehow "measuring the non uniqueness"), we deduce that in order to have non uniqueness, Φ_δ must blow up at least as fast as $\log(1/\delta)$ as $\delta \to 0$. More in general, we understand that upper bounds on Φ_δ are going to provide upper estimates on the measure of the superlevels $\{|\boldsymbol{X}_1 - \boldsymbol{X}_2| > \gamma\}$.

6.2 The maximal function

Since we have now dropped the Lipschitz context, the difference quotients

$$\frac{|b(x) - b(y)|}{|x - y|}$$

are no more uniformly bounded. In the proof of the upper bounds for the functional Φ_δ, we will have to deal again with difference quotients, but we will take advantage of the fact that they will always appear under the integral sign.

Suitable estimates can be provided using the classical maximal function. Given $f \in L^1_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^m)$, we define its maximal function as

$$Mf(x) = \sup_{r > 0} \fint_{B(x,r)} |f(y)| \, dy, \qquad x \in \mathbb{R}^d. \tag{47}$$

Similarly, when μ is a $\mathbb{R}^m$-valued measure in $\mathbb{R}^d$ with locally finite total variation, we define

$$M\mu(x) = \sup_{r > 0} \frac{|\mu|(B(x,r))}{\mathscr{L}^d(B(x,r))}, \qquad x \in \mathbb{R}^d. \tag{48}$$

Notice that since $|f\mathscr{L}^d| = |f|\mathscr{L}^d$, the two definitions are consistent, namely $M(f\mathscr{L}^d) = Mf$. Notice that, if $f \in L^1(\mathbb{R}^d; \mathbb{R}^m)$, or if μ is a measure with globally finite total variation in $\mathbb{R}^d$, then the maximal function is $\mathscr{L}^d$-a.e. finite.

A fundamental question is to provide bounds for this operator. It is immediate to see that

$$\|Mf\|_{L^\infty} \leq \|f\|_{L^\infty}. \tag{49}$$

Remarkably, the analogue property involving L^1 does *not* hold. Only the *weak estimate*

$$|||Mf|||_{M^1} \leq C_{d,1} \|f\|_{L^1} \tag{50}$$

holds, where the weak Lebesgue space $M^1(\mathbb{R}^d)$ is defined as the space consisting of all measurable function g on $\mathbb{R}^d$ such that

$$|||g|||_{M^1} = \sup_{\lambda > 0} \left\{ \lambda \mathscr{L}^d\big(\{x : |g(x)| < \lambda\}\big) \right\} < \infty.$$

Note carefully that the quantity $|||g|||_{M^1}$ is *not* a norm, and for this reason we have chosen the notation with the triple vertical bar.

By interpolating (49) and (50) we can obtain the *strong estimate*

$$\|Mf\|_{L^p} \leq C_{d,p} \|f\|_{L^p}, \tag{51}$$

valid for every $1 < p \leq \infty$, with a constant $C_{d,p}$ which blows up as $p \to 1$.

The maximal function is relevant in our context because of the following property. Assume that $f \in BV(\mathbb{R}^d)$. Then there exists a Lebesgue negligible set $N \subset \mathbb{R}^d$ such that

$$|f(x) - f(y)| \leq C|x - y|\Big(MDf(x) + MDf(y)\Big) \qquad \forall x, \, y \in \mathbb{R}^d \setminus N, \tag{52}$$

where C is a dimensional constant. Estimate (52) provides a control of the difference quotients with the maximal function of the derivative. We are going to exploit this property instead of the Lipschitz assumption in the estimates for Φ_δ.

Exercise 37 Prove formula (52) (a detailed proof is, for instance, in Theorem 5.34 of [9]).

For a detailed exposition of the theory of maximal functions we refer for instance to [98].

6.3   Upper bound for the integral quantity

We derive now an upper bound on Φ_δ which makes use of the dynamics of our situation, that is, we start to exploit the fact that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are regular Lagrangian flows. The following estimates will rely on the $W^{1,p}$ regularity of the vector fields $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ (we assume $p > 1$ in this section, while $p = 1$ will be considered in the next section), and on the compressibility constants of the regular Lagrangian flows. Also notice that, if $x \in B_R$, then $\boldsymbol{X}_i(t, x) \in B_{R+T\|\boldsymbol{b}_i\|_\infty}$ for all $t \in [0, T]$.

We estimate the time derivative of Φ_δ as follows:

$$
\begin{aligned}
\Phi_\delta'(t) &\le \int_{B_R} \frac{|\boldsymbol{b}_1(\boldsymbol{X}_1) - \boldsymbol{b}_2(\boldsymbol{X}_2)|}{\delta + |\boldsymbol{X}_1 - \boldsymbol{X}_2|}\, dx \\
&\le \int_{B_R} \frac{|\boldsymbol{b}_1(\boldsymbol{X}_2) - \boldsymbol{b}_2(\boldsymbol{X}_2)|}{\delta + |\boldsymbol{X}_1 - \boldsymbol{X}_2|}\, dx + \int_{B_R} \frac{|\boldsymbol{b}_1(\boldsymbol{X}_1) - \boldsymbol{b}_1(\boldsymbol{X}_2)|}{\delta + |\boldsymbol{X}_1 - \boldsymbol{X}_2|}\, dx \\
&\le \frac{1}{\delta}\int_{B_R} |\boldsymbol{b}_1(\boldsymbol{X}_2) - \boldsymbol{b}_2(\boldsymbol{X}_2)|\, dx + \int_{B_R} \min\left\{ \frac{2\|\boldsymbol{b}_1\|_\infty}{\delta} \; ; \; \frac{|\boldsymbol{b}_1(\boldsymbol{X}_1) - \boldsymbol{b}_1(\boldsymbol{X}_2)|}{|\boldsymbol{X}_1 - \boldsymbol{X}_2|} \right\}\, dx \\
&\le \frac{L_2}{\delta}\|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_{L^1\left(B_{R+T\|\boldsymbol{b}_2\|_\infty}\right)} + \int_{B_R} \min\left\{ \frac{2\|\boldsymbol{b}_1\|_\infty}{\delta} \; ; \; C\Big(M\nabla\boldsymbol{b}_1(\boldsymbol{X}_1) + M\nabla\boldsymbol{b}_1(\boldsymbol{X}_2) \Big) \right\}\, dx\,.
\end{aligned}
$$
$$\tag{53}$$

We can estimate

$$
\begin{aligned}
\int_{B_R} M\nabla\boldsymbol{b}_1(\boldsymbol{X}_i)\, dx \;&\le\; L_i \int_{B_{R+T\|\boldsymbol{b}_i\|_\infty}} M\nabla\boldsymbol{b}_1(x)\, dx \\
&\le\; C\|M\nabla\boldsymbol{b}_1\|_{L^p} \;\le\; C\|\nabla\boldsymbol{b}_1\|_{L^p}\,,
\end{aligned}
$$

where the constant C depends on L_i, p, $\|\boldsymbol{b}_i\|_\infty$, R and T. Note that with a bit more care this estimate could be made local, so to depend on the norm of $\nabla\boldsymbol{b}_1$ on a compact set only.
Hence, for a vector field $\boldsymbol{b}_1 \in W^{1,p}$ with $p > 1$ we deduce

$$
\Phi_\delta'(t) \le \frac{C}{\delta}\|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_{L^1_x} + C\|\nabla\boldsymbol{b}_1\|_{L^p_x}\,,
$$

and so, since $\Phi_\delta(0) = 0$, we have

$$
\Phi_\delta(t) \le \frac{C}{\delta}\|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_{L^1_t(L^1_x)} + C\|\nabla\boldsymbol{b}_1\|_{L^1_t(L^p_x)}\,.
$$

Putting this together with (46) we conclude

$$
\begin{aligned}
\mathscr{L}^d\Big(B_R &\cap \{|\boldsymbol{X}_1(t,\cdot) - \boldsymbol{X}_2(t,\cdot)| > \gamma\}\Big) \\
&\le \frac{C}{\delta \log\left(1 + \frac{\gamma}{\delta}\right)}\|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_{L^1_t(L^1_x)} + \frac{C}{\log\left(1 + \frac{\gamma}{\delta}\right)}\|\nabla\boldsymbol{b}_1\|_{L^1_t(L^p_x)}\,.
\end{aligned}
$$
$$\tag{54}$$

This is the fundamental estimate, which will allow us to deduce many of the well-posedness results for the ODE.

6.4   Well-posedness of the ODE: $W^{1,p}$ regularity

We start with *uniqueness*, which is an immediate consequence of (54). Since we look at $\boldsymbol{b}_1 = \boldsymbol{b}_2$, we have

$$
\mathscr{L}^d\Big(B_R \cap \{|\boldsymbol{X}_1(t,\cdot) - \boldsymbol{X}_2(t,\cdot)| > \gamma\}\Big) \le \frac{C}{\log\left(1 + \frac{\gamma}{\delta}\right)}\|\nabla\boldsymbol{b}_1\|_{L^1_t(L^p_x)}\,,
$$

for every $\delta > 0$ and every $R > 0$. It suffices to let $\delta \to 0$.

Next, we discuss *stability*. Consider $\boldsymbol{b} \in L_t^1(W_x^{1,p})$ and a sequence $(\boldsymbol{b}_n)$ convergent to $\boldsymbol{b}$ in L_{loc}^1, equibounded in L^∞. Assume that the regular Lagrangian flows $\boldsymbol{X}$ and $\boldsymbol{X}_n$ have equibounded compressibility constants. The fundamental estimate (54) now reads

$$
\mathscr{L}^d\Big(B_R \cap \{|\boldsymbol{X}_n(t,\cdot) - \boldsymbol{X}(t,\cdot)| > \gamma\}\Big)
$$
$$
\leq \frac{C}{\delta \log\left(1 + \dfrac{\gamma}{\delta}\right)} \|\boldsymbol{b}_n - \boldsymbol{b}\|_{L_t^1(L_{x,\mathrm{loc}}^1)} + \frac{C}{\log\left(1 + \dfrac{\gamma}{\delta}\right)} \|\nabla \boldsymbol{b}\|_{L_t^1(L_x^p)} \; = \; I + II\,.
$$

Given $\gamma, \eta > 0$, we choose $\delta > 0$ so small that $II \leq \eta/2$. This fixes the quantity $\frac{C}{\delta \log(1+\gamma/\delta)}$ in I. Therefore we can find $\bar{n}$ so large that $I \leq \eta/2$ for all $n \geq \bar{n}$.

Hence we have discovered that: given $\gamma > 0$ and $R > 0$, for every $\eta > 0$ we can find $\bar{n}$ such that

$$
\mathscr{L}^d\Big(B_R \cap \{|\boldsymbol{X}_n(t,\cdot) - \boldsymbol{X}(t,\cdot)| > \gamma\}\Big) \leq \eta \qquad \forall n \geq \bar{n}\,.
$$

This means that $\boldsymbol{X}_n$ is converging to $\boldsymbol{X}$ locally in measure in $\mathbb{R}^d$. Since $\boldsymbol{X}$ and $\boldsymbol{X}_n$ are locally equibounded, this also implies L_{loc}^1 convergence: given $R > 0$, for every $\eta > 0$ we can find $\bar{n}$ such that

$$
\int_{B_R} |\boldsymbol{X}_n(t,x) - \boldsymbol{X}(t,x)|\, dx \leq \eta \qquad \forall n \geq \bar{n}\,.
$$

Note the important fact that stability comes with a quantitative rate: the values $\bar{n}$ only depends on η and on the equibounds on the sequence.

The proof of the *compactness* goes along the same line. Consider a sequence $(\boldsymbol{b}_n)$ which is equibounded in L^∞ and in $L_t^1(W_x^{1,p})$. Assume that there exist associated regular Lagrangian flows $\boldsymbol{X}_n$ with equibounded compressibility constants. Then we can use the fundamental estimate (54) to write

$$
\mathscr{L}^d\Big(B_R \cap \{|\boldsymbol{X}_n(t,\cdot) - \boldsymbol{X}_m(t,\cdot)| > \gamma\}\Big)
$$
$$
\leq \frac{C}{\delta \log\left(1 + \dfrac{\gamma}{\delta}\right)} \|\boldsymbol{b}_n - \boldsymbol{b}_m\|_{L_t^1(L_x^1)} + \frac{C}{\log\left(1 + \dfrac{\gamma}{\delta}\right)} \; = \; I + II\,,
$$

where the constant C in II also depends on the equibounds on $\|D\boldsymbol{b}_n\|_{L_t^1(L_x^p)}$.

Given $\eta > 0$, we find $\delta > 0$ such that $II \leq \eta/2$. Correspondingly, we find $\bar{n}$ such that $I \leq \eta/2$ for every $n, m \geq \bar{n}$.

Summarizing, we have proven that, given $\gamma > 0$ and $R > 0$, for every $\eta > 0$ we can find $\bar{n}$ such that

$$
\int_{B_R} |\boldsymbol{X}_n(t,x) - \boldsymbol{X}_m(t,x)|\, dx \leq \eta \qquad \forall n, m \geq \bar{n}\,,
$$

that is, the sequence $(\boldsymbol{X}_n)$ is locally precompact in measure. As for the stability, thanks to the local equiboundedness of $\boldsymbol{X}_n$, precompactness in L_{loc}^1 can be deduced from this.

Remark 38 The analogue compactness statement under L^1 bounds on $\nabla\boldsymbol{b}_n$, instead of L^p bounds, is the content of a conjecture due to Bressan [45]. The $W^{1,1}$ theory presented in the next section will replace the L^p bounds on $\nabla\boldsymbol{b}_n$ with equi-integrability bounds.

The compactness is the key property in order to prove **existence** of the regular Lagrangian flow associated to a bounded vector field $\boldsymbol{b} \in L^1_t(W^{1,p}_x)$. At this point, we have to assume some condition about the divergence of $\boldsymbol{b}$ and, for simplicity, we assume div $\boldsymbol{b} \in L^\infty$.
We regularize $\boldsymbol{b}$ by convolution, in order to get a sequence $\boldsymbol{b}_n$ of smooth vector fields which are equibounded in $L^\infty \cap L^1_t(W^{1,p}_x)$, and with div $\boldsymbol{b}_n$ equibounded in L^∞. Remembering (7), we infer that the (classical) flows associated to $\boldsymbol{b}_n$ have equibounded compressibility constant. We are therefore in the position to apply the compactness property, and it is simple to prove that this produces a regular Lagrangian flow associated to $\boldsymbol{b}$.

6.5   A mild regularity property of the regular Lagrangian flow

Remember that, for Lipschitz vector fields, the flow inherits regularity with respect to the spatial variable: $x \mapsto \boldsymbol{X}(t,x)$ is Lipschitz, with Lipschitz constant depending exponentially on the time and on the Lipschitz constant of the vector field (compare once again (2)).
We want now to derive a similar result for Sobolev vector fields. In order to do this, we consider a functional Ψ_δ, strictly related to the functional $\Phi_\delta(t)$ in (45), that is

$$\Psi(\boldsymbol{X}) = \left\| \sup_{0\leq t \leq T} \sup_{r>0} \fint_{B(x,r)} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r}\right) dy \right\|_{L^p_x(B_R)}, \tag{55}$$

where $R > 0$ is fixed.
We want to derive upper bounds for the functional $\Psi(\boldsymbol{X})$. We begin by differentiating with respect to time the integral:

$$\frac{d}{dt} \fint_{B(x,r)} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r}\right) dy$$

$$\leq \fint_{B(x,r)} \frac{|\boldsymbol{b}(\boldsymbol{X}(t,x)) - \boldsymbol{b}(\boldsymbol{X}(t,y))|}{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|} dy$$

$$\leq C \fint_{B(x,r)} \left[M\nabla\boldsymbol{b}(\boldsymbol{X}(t,x)) + M\nabla\boldsymbol{b}(\boldsymbol{X}(t,y)) \right] dy$$

$$= CM\nabla\boldsymbol{b}(\boldsymbol{X}(t,x)) + C \fint_{B(x,r)} M\nabla\boldsymbol{b}(\boldsymbol{X}(t,y)) \, dy \,.$$

Integrating in time we deduce

$$\fint_{B(x,r)} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r}\right) dy \leq \fint_{B(x,r)} \log\left(1 + \frac{|x-y|}{r}\right) dy$$

$$+ C \int_0^t M\nabla\boldsymbol{b}(\boldsymbol{X}(s,x)) \, ds + C \int_0^t \fint_{B(x,r)} M\nabla\boldsymbol{b}(\boldsymbol{X}(s,y)) \, dy \, ds$$

$$\leq \log 2 + C \int_0^t M\nabla\boldsymbol{b}(\boldsymbol{X}(s,x)) \, ds + C \int_0^t \fint_{B(x,r)} M\nabla\boldsymbol{b}(\boldsymbol{X}(s,y)) \, dy \, ds \,.$$

Taking the sup over $0 \leq t \leq T$ and over $r > 0$, and finally the $L_x^p(B_R)$ norm we deduce

$$
\begin{aligned}
\Psi(\boldsymbol{X}) &\leq C + C \int_0^T \|M \nabla \boldsymbol{b}(\boldsymbol{X}(s,x))\|_{L^p(B_R)}\, ds + C \int_0^T \left\| M\big[M\nabla \boldsymbol{b}(\boldsymbol{X}(s,\cdot))\big] \right\|_{L^p(B_R)} ds \\
&\leq C + C \int_0^T \|M \nabla \boldsymbol{b}\|_{L^p}\, ds + C \int_0^T \|M\nabla \boldsymbol{b}(\boldsymbol{X}(s,\cdot))\|_{L^p}\, ds \\
&\leq C + C\|\nabla \boldsymbol{b}\|_{L_t^1(L_x^p)} + C \int_0^T \|M \nabla \boldsymbol{b}\|_{L^p}\, ds \\
&\leq C + C\|\nabla \boldsymbol{b}\|_{L_t^1(L_x^p)} \leq C\,,
\end{aligned}
\tag{56}
$$

where the constant C only depends on the bounds on $\boldsymbol{b}$ and on the compressibility constant of the regular Lagrangian flow $\boldsymbol{X}$.

We are going to utilize Chebyshev inequality, which asserts that for every $1 \leq p < \infty$ and every $\lambda > 0$ we can estimate

$$
\mathscr{L}^d\big(\{|f| > \lambda\}\big) \leq \frac{\|f\|_{L^p}^p}{\lambda^p}\,.
$$

Applying this with $\lambda = C/\epsilon^{1/p}$ to the function of x introduced in (55), we obtain the existence of a set $K \subset B_R$ with $\mathscr{L}^d(B_R \setminus K) \leq \epsilon$ such that

$$
\sup_{0 \leq t \leq T} \sup_{r > 0} \fint_{B(x,r)} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r}\right) dy \leq \frac{C}{\epsilon^{1/p}} \qquad \forall x \in K\,.
$$

This means that

$$
\fint_{B(x,r)} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r}\right) dy \leq \frac{C}{\epsilon^{1/p}}
$$

for every $x \in K$, every $r > 0$ and every $0 \leq t \leq T$.

We apply this in order to derive the Lipschitz estimate for the regular Lagrangian flow $\boldsymbol{X}$. Let x and x' be points of K, and set $r = |x - x'|$. Define

$$
C_{x,x'} = B(x,r) \cap B(x',r)\,.
$$

Notice that $\mathscr{L}^d(C_{x,x'}) = c_d r^d$, for a positive dimensional constant c_d. We can estimate

$$
\begin{aligned}
\log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,x')|}{r}\right) &= \fint_{C_{x,x'}} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,x')|}{r}\right) dy \\
&\leq \fint_{C_{x,x'}} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r}\right) dy + \fint_{C_{x,x'}} \log\left(1 + \frac{|\boldsymbol{X}(t,x') - \boldsymbol{X}(t,y)|}{r}\right) dy \\
&\leq C \fint_{B(x,r)} \log\left(1 + \frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r}\right) dy + C \fint_{B(x',r)} \log\left(1 + \frac{|\boldsymbol{X}(t,x') - \boldsymbol{X}(t,y)|}{r}\right) dy \\
&\leq \frac{C}{\epsilon^{1/p}}\,.
\end{aligned}
$$

As a consequence

$$|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,x')| \leq |x - x'| \exp\left(\frac{C}{\epsilon^{1/p}}\right)$$

for every x and x' belonging to K, that is

$$\operatorname{Lip} \boldsymbol{X}(t,\cdot)|_K \leq \exp\left(\frac{C}{\epsilon^{1/p}}\right). \tag{57}$$

Remember that the set $K \subset B_R$ was such that $\mathscr{L}^d(B_R \setminus K) \leq \epsilon$.

Remark 39 The Lipschitz property we have proven is usually called Lusin-type approximation with Lipschitz function. From this, approximate differentiability $\mathscr{L}^d$-a.e. of the map $x \mapsto \boldsymbol{X}(t,x)$ can be deduced.

Exercise 40 Using Ascoli-Arzelá Theorem, find another proof of the compactness of the regular Lagrangian flow. Notice that it is essential that the control on the Lipschitz constant provided by (57) is quantitative.

7 Quantitative ODE estimates: $W^{1,1}$ regularity and singular integrals

In the previous section we have described the quantitative estimates and their consequences in the case of $W^{1,p}$ regularity of the vector field, when $p > 1$. The case $p = 1$ (and a fortiori the BV case) was not covered by this analysis, due to the lack of strong estimates for $p = 1$ for the maximal function: (51) does not hold for $p = 1$, and only the weak estimate (50) is available. However, (50) seems to be of no help for our purposes: we need to bound an integral of difference quotients, which are in turn bounded by the maximal function of the gradient of the vector field. The bound in M^1 coming from (50) does not even guarantee integrability!

The good news is that there is an additional element that can be exploited in our analysis, and this refined proof will allow to obtain the same theory (except for the regularity results in Section 6.5) in the $W^{1,1}$ case (but unfortunately not in the BV case, presently only covered by the theory of renormalized solutions).

In the derivation of the upper bound for the functional $\Phi_\delta(t)$, after differentiation in time, we obtained the bound

$$\Phi'_\delta(t) \leq \frac{L_2}{\delta}\|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_{L^1} + \int_{B_R} \min\left\{\frac{2\|b_1\|_\infty}{\delta} \; ; \; C\left(M\nabla\boldsymbol{b}_1(\boldsymbol{X}_1) + M\nabla\boldsymbol{b}_1(\boldsymbol{X}_2)\right)\right\} dx.$$

Afterwards, we proceeded with the estimate of the difference quotients, simply "forgetting" the first term inside the minimum.

In this section we describe how an interpolation argument (exploiting the L^∞ term $\frac{2\|\boldsymbol{b}\|_\infty}{\delta}$ in the above minimum) will allow to recover both the $W^{1,1}$ case, and the case in which the derivative of the vector field is a singular integral of an L^1 function. Notice that this second class is not included in, nor it includes, BV. We also remark that equi-integrability bounds will play a role in the estimates: this is the obstruction in pushing this $W^{1,1}$ technique to BV.

7.1 Weak Lebesgue spaces and interpolation

In order to address the case $p = 1$, some additional tools will be needed. Recall that the weak Lebesgue space $M^1(\mathbb{R}^d)$ is defined as the space consisting of all measurable function g on $\mathbb{R}^d$ such that

$$|||g|||_{M^1} = \sup_{\lambda > 0} \left\{ \lambda \mathscr{L}^d \big(\{ x : |g(x)| < \lambda \} \big) \right\} < \infty \,.$$

We already noticed that the maximal function satisfies a weak inequality from L^1 into M^1, expressed by (50). The same holds when we consider the maximal function of a measure, defined as in (48):

$$|||M\mu|||_{M^1} \le C_{d,1} \|\mu\|_{\mathcal{M}} \,, \tag{58}$$

where on the right hand side we have the total variation norm.

The space M^1 is strictly bigger than L^1. For instance, on the real line, $1/x$ belongs to M^1 but fails to be in L^1 (even locally, due to the blow up at the origin). The following interpolation inequality may be interpreted by informally saying that "M^1 is not too distant from L^1":

$$\|f\|_{L^1} \le |||f|||_{M^1} \left[1 + \log \left(\frac{C\|f\|_{L^\infty}}{|||f|||_{M^1}} \right) \right] \,. \tag{59}$$

Indeed, we can bound the L^1 norm with the M^1 (pseudo)norm, up to a logarithm of the L^∞ norm. We notice that (59) holds for functions defined on a compact set, and that one could similarly interpolate between M^1 and L^p, for any $p > 1$.

Exercise 41 Prove the interpolation inequality (59). In order to do this, rewrite the L^1 norm of f as the integral of the distribution function of f defined as usual by $\alpha(\lambda) = \mathscr{L}^d(\{|f| > \lambda\})$. Split the integral in the sum of the integrals over $[0, \Lambda]$ and $[\Lambda, \|f\|_{L^\infty}]$, use suitable estimates in the two intervals, and conclude by optimizing over Λ.

7.2 Uniqueness for the ODE: $W^{1,1}$ regularity

We show now how to extend the estimates in Section 6.3 to include the case $p = 1$. For clearness and simplicity of exposition, we focus our presentation on the uniqueness issue, but all the well-posedness results of Section 6.4 could be proven in this context along the same lines. However, it is less clear how to extend to this case the Lischitz estimates in Section 6.5.

Given a vector field $\boldsymbol{b} \in W^{1,1}$, we consider two regular Lagrangian flows $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, with compressibility constants given respectively by L_1 and L_2. Proceeding as in (53) we estimate

$$\Phi_\delta'(t) \le \int_{B_R} \min \left\{ \frac{2\|\boldsymbol{b}\|_\infty}{\delta} \,;\, C\Big(MD\boldsymbol{b}(\boldsymbol{X}_1) + MD\boldsymbol{b}(\boldsymbol{X}_2) \Big) \right\} dx$$

$$\le C(L_1 + L_2) \int_{B_{R+T\|\boldsymbol{b}\|_\infty}} \min \left\{ \frac{\|\boldsymbol{b}\|_\infty}{\delta} \,;\, MD\boldsymbol{b} \right\} dx \,.$$

By integrating with respect to the time we arrive at

$$\Phi_\delta(t) \leq C(L_1 + L_2) \int_0^T \int_{B_{R+T\|\boldsymbol{b}\|_\infty}} \min\left\{\frac{\|\boldsymbol{b}\|_\infty}{\delta} \ ; \ MD\boldsymbol{b}\right\} dx\, ds\,. \tag{60}$$

We now set

$$\phi(s, x) = \min\left\{\frac{\|\boldsymbol{b}\|_\infty}{\delta} \ ; \ MD\boldsymbol{b}\right\} \tag{61}$$

and observe that we have the two bounds

$$\|\phi\|_{L_{tx}^\infty} \leq \frac{\|\boldsymbol{b}\|_\infty}{\delta}\,, \qquad |||\phi|||_{M_{tx}^1} \leq |||MD\boldsymbol{b}|||_{M_{tx}^1}\,. \tag{62}$$

We can exploit the interpolation inequality (59) and the bounds (62), together with the fact that the function $z \mapsto \log z$ is increasing and the function $z \mapsto z[1 + \log(c/z)]$ is increasing for $z \leq c/\delta$, to obtain

$$
\begin{aligned}
\Phi_\delta(t) &\leq C|||\phi|||_{M_{tx}^1}\left[1 + \log\left(\frac{C\|\phi\|_{L_{tx}^\infty}}{|||\phi|||_{M_{tx}^1}}\right)\right] \\
&\leq C|||MD\boldsymbol{b}|||_{M_{tx}^1}\left[1 + \log\left(\frac{C\|\boldsymbol{b}\|_{L_{tx}^\infty}}{\delta|||MD\boldsymbol{b}|||_{M_{tx}^1}}\right)\right]\,.
\end{aligned}
$$

From (58) we simply obtain

$$|||MD\boldsymbol{b}|||_{M_{tx}^1} \leq \left\||||MD\boldsymbol{b}|||_{M_x^1}\right\|_{L_t^1} \leq C\|D\boldsymbol{b}\|_{L_t^1(\mathcal{M}_x)}\,,$$

and using again the fact that the function $z \mapsto z[1 + \log(c/z)]$ is increasing for $z \leq c/\delta$ this gives

$$\Phi_\delta(t) \leq C\|D\boldsymbol{b}\|_{L_t^1(\mathcal{M}_x)}\left[1 + \log\left(\frac{C}{\delta\|D\boldsymbol{b}\|_{L_t^1(\mathcal{M}_x)}}\right)\right]\,, \tag{63}$$

in which $\|\boldsymbol{b}\|_\infty$ has been absorbed in the constant.

This upper bound is exactly on the critical scale discriminating for uniqueness in (46)! This means that (63) (valid even for BV vector fields) is not enough to conclude.

However, in the $W^{1,1}$ case we can "play with the constants". Indeed, given $\epsilon > 0$, the derivative $\nabla\boldsymbol{b} \in L^1([0,T] \times \mathbb{R}^d)$ can be split as

$$\nabla\boldsymbol{b} = g_\epsilon^1 + g_\epsilon^2\,, \tag{64}$$

with $\|g_\epsilon^1\|_{L_t^1(L_x^1)} \leq \epsilon$ and $\|g_\epsilon^2\|_{L_t^1(L_x^2)} \leq C_\epsilon$. The constant C_ϵ blows up as $\epsilon \to 0$ and depends on the equi-integrability of $\nabla\boldsymbol{b}$. Using this argument, we are able to proceed in the $W^{1,1}$ case only, and the more general BV case (still allowed in (63)) has to be abandoned. A decomposition as in (64) does not hold for a measure.

With this decomposition we can estimate the function ϕ in (61) as follows:

$$
\begin{aligned}
\phi(s,x) &\leq \min\left\{\frac{\|\boldsymbol{b}\|_\infty}{\delta}\;;\;Mg_\epsilon^1 + Mg_\epsilon^2\right\} \\
&\leq \min\left\{\frac{\|\boldsymbol{b}\|_\infty}{\delta}\;;\;Mg_\epsilon^1\right\} + \min\left\{\frac{\|\boldsymbol{b}\|_\infty}{\delta}\;;\;Mg_\epsilon^2\right\} = \phi^1(s,x) + \phi^2(s,x)\,.
\end{aligned}
$$

Reasoning as before and exploiting (64) we have

$$
\|\phi^1\|_{L_{tx}^\infty} \leq \frac{\|\boldsymbol{b}\|_\infty}{\delta}\,,\qquad |||\phi^1|||_{M_{tx}^1} \leq |||Mg_\epsilon^1|||_{M_{tx}^1} \leq \left\||||Mg_\epsilon^1|||_{M_x^1}\right\|_{L_t^1} \leq C\|g_\epsilon^1\|_{L_t^1(L_x^1)} \leq C\epsilon
$$

and

$$
\|\phi_\epsilon^2\|_{L_t^1(L_x^2)} \leq C\|g_\epsilon^2\|_{L_t^1(L_x^2)} \leq CC_\epsilon\,.
$$

We use these two estimates in the bound for Φ_δ. The "borderline" estimate (63) is utilized with the term g_ϵ^1, while the term g_ϵ^2 is treated according to the analysis for the case $p > 1$ in Section 6.3. We finally get an estimate of the form

$$
\mathscr{L}^d\big(B_R \cap \{|\boldsymbol{X}_1 - \boldsymbol{X}_2| > \gamma\}\big) \leq C\epsilon\,\frac{1 + \log\left(\frac{C}{\delta\epsilon}\right)}{\log\left(1 + \frac{\gamma}{\delta}\right)} + \frac{CC_\epsilon}{\log\left(1 + \frac{\gamma}{\delta}\right)}\,,
$$

and we conclude by choosing first ϵ and then δ sufficiently small.

Remark 42 The extension of this proof to the BV case is an important open problem. This would solve the compactness conjecture proposed by Bressan in [45].

7.3   An extension to a case involving singular integrals

We want to extend the argument in Section 7.2 to a more general case, involving singular integrals of L^1 functions. We start by presenting a motivation for this setting.
The Euler equation in vorticity form in two dimension reads

$$
\partial_t \omega + \operatorname{div}(\boldsymbol{v}\,\omega) = 0\,. \tag{65}
$$

This is a continuity equation for the fluid vorticity ω, which is transported by the (divergence free) fluid velocity $\boldsymbol{v}$. This is indeed a *nonlinear* continuity equation, due to the structural condition $\omega = \operatorname{curl}\boldsymbol{v}$. Both the vorticity and the velocity are unknowns of the problem, and the coupling can be rewritten as a convolution:

$$
\boldsymbol{v} = K * \omega\,. \tag{66}
$$

In the above formula, the Biot-Savart kernel is defined (up to multiplicative constants) by $K(x) = x^\perp/|x|^2$.
In order to proceed with the analysis of (65), it is necessary to understand the regularity of the fluid velocity $\boldsymbol{v}$. A formal differentiation if the convolution product in (66) gives $D\boldsymbol{v} = DK * \omega$.

The problem is now that $|DK| \sim 1/|x|^2$ around the origin in $\mathbb{R}^2$, hence it is in general not even locally integrable. We cannot give a meaning to the formula for the derivative of $\boldsymbol{v}$ as a usual convolution of distributions.

However, the actual expression of DK (without the modulus!) retains enough cancellations to allow for a definition of the "convolution" in the sense of Calderón-Zygmund singular integrals (see [98] for a detailed description of this topic).

The analysis of Section 7.2 extends naturally to the case in which the derivative of the vector field $\boldsymbol{b}$ is given by singular integrals of L^1 functions:

$$\partial_j \boldsymbol{b}^i = \sum_k S_{ijk}\, g_{ijk}\,,$$

where $g_{ijk} \in L^1$ and each S_{ijk} is a singular integral operator, associated to a kernel $K_{ijk} \in \mathcal{S}'(\mathbb{R}^d) \cap C^1(\mathbb{R}^d \setminus \{0\})$, satisfying

$$|K(x)| \le \frac{C}{|x|^d}\,, \qquad |DK(x)| \le \frac{C}{|x|^{d+1}}$$

and the cancellation property

$$\left| \int_{R_1 < |x| < R_2} K(x)\, dx \right| \le C \qquad \forall\, 0 < R_1 < R_2 < \infty\,.$$

Remark 43 This class of vector fields is motivated by the example of the two-dimensional Euler equation with L^1 vorticity. It includes $W^{1,1}$, but is strictly bigger; on the other hand, it is not included in, nor it includes, BV.

We now describe the main ideas from this proof. Analogously to maximal functions, singular integrals satisfy strong estimates

$$\|Sf\|_{L^p} \le C_p \|f\|_{L^p}$$

for every $1 < p < \infty$, but only the weak estimate

$$|||Sf|||_{M^1} \le C_1 \|f\|_{L^1}$$

when $p = 1$. Given two singular integral operators S_1 and S_2, associated to regular enough singular kernel K_1 and K_2, the composition $S = S_2 \circ S_1$ is again a singular integral operator, associated to the singular kernel $K = K_2 * K_1$. This means that the weak estimate

$$|||Sf|||_{M^1} = |||S_2 \circ S_1 f|||_{M^1} \le C_1 \|f\|_{L^1}$$

holds. Notice carefully that this estimate cannot be obtained by composing the two weak estimates from L^1 into M^1 valid for S_1 and S_2 separately. The idea is that *cancellations* in the formal convolution product $K_2 * K_1$ are exploited.

We go back to our uniqueness argument, involving the search for upper bounds for the functional Φ_δ. We would like to obtain something similar to (63) (with $\|g\|_{L^1}$ instead of $\|D\boldsymbol{b}\|_{L^1}$ on the right hand side). We realize that we need an estimate of the form

$$|||MD\boldsymbol{b}|||_{M^1} = |||MSg|||_{M^1} \le C\|g\|_{L^1} . \tag{67}$$

The classical maximal function in (47) is "too rough" to allow for such an estimate. The idea is that we can choose a "smooth maximal function", with absolute value outside the integral, and with a smooth weight. More precisely, given $\rho \in C_c^\infty$, we define

$$M_\rho f(x) = \sup_{r>0} \left| \frac{1}{r^d} \int_{\mathbb{R}^d} \rho \left(\frac{x-y}{r} \right) f(y) \, dy \right| .$$

It can be proven that this smooth maximal function (also known as the grand maximal function in the context of Hardy spaces) is still suitable for the estimate of difference quotients as in (52). Moreover, the cancellations properties of the kernel ρ play together with those of the singular integral S, allowing to get estimate (67). We refer to [36] for the exact statements and for the proofs.

8   Some applications

8.1   A system of conservation laws.

Let us consider the Cauchy problem (studied in one space dimension by Keyfitz–Kranzer in [82])

$$\frac{d}{dt}u + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \left(\boldsymbol{f}_i(|u|)u \right) = 0, \qquad u : \mathbb{R}^d \times (0, +\infty) \to \mathbb{R}^k \tag{68}$$

with the initial condition $u(\cdot, 0) = \bar{u}$. Here $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^d$ is a C^1 function.
In [46], 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 [11] it has been proved that this procedure can really be implemented, thanks to the results in [10], 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 [14] it has been proved that the lower bound on $\bar{\rho}$ is not necessary and, moreover, that the solution built in [11] is unique in a suitable class of admissible functions: those whose modulus ρ satisfies the scalar PDE

$$\frac{d}{dt}\rho + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \left(\boldsymbol{f}_i(\rho)\rho \right) = 0 \tag{69}$$

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 $(s\boldsymbol{f})'(s)\eta'(s) = \boldsymbol{q}'(s)$, see [65] for existence, uniqueness and regularity results for Kruzhkov solutions), with the initial condition $\rho(0, \cdot) = \bar{\rho}$.

Notice that the regularity theory for this class of solutions gives that $\rho \in L^\infty \cap BV_{\text{loc}}\left([0, +\infty) \times \mathbb{R}^d\right)$, due to the BV regularity and the boundedness of $|\bar{u}|$. Furthermore the maximum principle gives $0 < 1/\rho \leq 1/|\bar{u}| \in L^\infty$.

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

$$\theta_t + \sum_{i=1}^{d} \frac{\partial}{\partial x_i}\left(\boldsymbol{f}_i(\rho)\theta\right) = 0 \tag{70}$$

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 (69).

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 (69), that could in principle 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 [11] for details).

8.2 Lagrangian solutions of semi-geostrophic equations.

The semigeostrophic equations are a simplifies model of the atmosphere/ocean flows [62], described by the system of transport equations

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

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

$$\frac{d}{dt}\nabla P + \boldsymbol{u} \cdot \nabla^2 P = J(\nabla P - x) \quad \text{with} \quad J := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{71}$$

Existence of solutions has been for this problem has been open until very recent times, while uniqueness is still open. In [28, 63], 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

$$\frac{d}{dt}\alpha_t + D_x \cdot (\boldsymbol{U}_t \alpha_t) = 0 \quad \text{with} \quad \boldsymbol{U}_t(X) := J(X - \nabla P_t^*(X)) \tag{72}$$

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 (72) 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 [64], 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 $\mathscr{L}_\Omega^d$.

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 (71), i.e.

$$\frac{d}{dt}Z_t(x) = J(Z_t - F_t) \tag{73}$$

This provides us with a sort of weak solution of (71). In connection with existence of solutions to the semi-geostrophic problem in the original physical variables, a formal argument suggests that, given P_t, the velocity $\boldsymbol{u}$ should be defined by

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

On the other hand, the a-priori regularity on ∇P_t (ensured by the convexity of P_t) is a BV regularity. In more recent times, using refined regularity results on Alexandrov solutions to Monge-Ampére equations which provide not only BV regularity, but also Sobolev regularity ([69, 70, 96], improving earlier results in [47, 48, 49, 50, 101, 102] in the borderline case when the right hand side is bounded away from 0 and ∞, but only Borel), in [23, 24] distributional solutions for the Eulerian form of the PDE have been recovered.

9 Open problems, bibliographical notes, and references

Section 2. The material contained in this section is classical. Good references are [75], Chapter 8 of [16], [42] and [72]. For the proof of the area formula, see for instance [9], [74], [78].
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 [16]. The same proof works under the weaker assumption (8).

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 [10], 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 [12], see also [13] which provides more insight into the quantitative side of the theory and the differentiability properties of the flow.
The idea of a probabilistic representation is of course classical, and appears in many contexts (particularly for equations of diffusion type); to our knowledge the first reference in the context of conservation laws and fluid mechanics is [37], where a similar approach is proposed for the incompressible Euler equation (see also [38, 39, 40]): 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 [95], and in the so called irrigation problems, a nice variant of the optimal transport problem [30]. See also [26] for a similar approach within Mather's theory. The Lecture Notes [105] (see also the Appendix of [88]) 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 [29]). Another related reference is [67].
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 [16], see also [89] 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 [97, 26].

In [20] an interesting variant of the stability Theorems 21 and 33 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 [20] that weak 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 [16] for a general proof of this fact, and for references to earlier works of Otto, Benamou-Brenier).

Section 4. The definition of renormalized solution and the strong convergence of commutators are entirely borrowed from [72]. See also [73] for the relevance of this concept in connection with the existence theory for Boltzmann equation.

The differentiability properties of the flow have been found in [83]: later on, this differentiability property has been characterized and compared with the more classical approximate differentiability ([78]) in [18], while [17] contains the proof of the stronger "local" Lipschitz properties. Theorem 34 summarizes all these results. The paper [61] contains also more explicit Lipschitz estimates and an independent proof of the compactness of flows.

See also [51] for a proof, using radial convolution kernels, of the renormalization property for vector fields satisfying $D_i \boldsymbol{b}^j + D_j \boldsymbol{b}^i \in L_{\mathrm{loc}}^1$.

Both methods, the one illustrated in these lecture notes and the DiPerna–Lions one, are based on abstract compactness arguments and do not provide a rate of convergence in the stability theorem.

A few existence result for Sobolev vector fields seem to be known in the infinite-dimensional case, see [31] and the more recent paper [22]. 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 $\{\tilde{\boldsymbol{b}} \ne \boldsymbol{b}\}$ 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 [10]. 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 levels of generality in [43, 85, 56] (see also [52, 53] for related results independent of this technique), until the general result [10].

The optimal regularity condition on $\boldsymbol{b}$ ensuring the renormalization property, and therefore the well-posedness in $\mathscr{L}_{\boldsymbol{b}}$, is still not known. New results, both in the Sobolev and in the *BV* framework, are presented in [15, 83, 84].

As mentioned in Remark 26, it is proved in [34] that the renormalization property can be *characterised* in terms of the uniqueness (for both the forward and the backward Cauchy problems) and the strong continuity of weak solutions, or in terms of the density of smooth functions with

respect to a suitable "graph norm". This is some sense tells that the renormalization property is more that a "technical tool" to prove uniqueness, but really a substantial property related to the well-posedness.

In [19] the possibility to prove the renormalization property for nearly incompressible $BV_{\mathrm{loc}} \cap L^\infty$ fields $\boldsymbol{b}$ is investigated: nearly incompressible fields are defined by requiring the existence of 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 [14], 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.

The situation in the two-dimensional context is somehow more "rigid". In the series of papers [5, 6, 7] a *characterisation* of the uniqueness for the continuity equation is proved, in the context of two-dimensional bounded autonomous divergence-free vector fields. To any such vector field b a Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}$ so that $b = \nabla^\perp f$ can be associated. The main result of [5] is the equivalence between the uniqueness and a "measure theoretical version" of the Sard property for the function f, which apart from a minor technical detail reads

$$ f_\#\big(\mathscr{L}^2 \llcorner \{\nabla f = 0\}\big) \perp \mathscr{L}^1 . \tag{74} $$

The heart of the proof is a suitable disintegration of the PDE along level sets of f. In this way, two-dimensional uniqueness is linked to uniqueness on the level sets. It turns out that (74) is the condition characterising uniqueness for such one-dimensional problems.

Section 6. The presentation in this section is based on [61]. Some related analysis, with further developments in the case of kinetic equations, is contained in [81, 54].

Section 7. Here the analysis follows [36], but in order to make the presentation more accessible, we skipped most of the details and we essentially tried to convey the general ideas. A similar presentation can be found in [35, 59].

The main open problem in this context is clearly the extension of the quantitative estimates to BV vector fields, or even more to vector fields whose gradient is given by singular integrals of measures (and not just of L^1 functions). In the work in progress [33] the following split-case is addressed: the derivatives with respect to some coordinates are singular integrals of measures, while all other derivatives are singular integrals of L^1 functions. The strategy is based on an anisotropic version of the functional.

The quantitative estimates in the context of singular integrals are relevant for applications to nonlinear PDEs. In particular, in [32] we plan to study existence and energy concentration for solutions to the two-dimensional Euler equation with L^1 vorticity, and some existence settings for the Vlasov-Poisson equation.

Section 8. 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 [14]). A strictly related problem is the convergence of the vanishing viscosity method to the solution built in [11]. Also, very little about the regularity of solutions is presently known: we know [66] that

BV estimates do not hold and, besides, that the construction in [11] seems not applicable to more general systems of triangular type, see the counterexample in [60].

Acknowledgement. This lecture notes have been written on the occasion of a Crash Course given by the authors at the Heriot-Watt University in Edinburgh on April 15-16, 2013. The authors warmly thank the organisers for the invitation and for the kind hospitality. They also thank Anna Bohun for a careful reading of a preliminary version of these notes.

References

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

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

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

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

[5] G. ALBERTI, S. BIANCHINI & G. CRIPPA: *A uniqueness result for the continuity equation in two dimensions.* Journal of the European Mathematical Society (JEMS), in press.

[6] G. ALBERTI, S. BIANCHINI & G. CRIPPA: *Structure of level sets and Sard-type properties of Lipschitz maps: results and counterexamples.* Annali della Scuola Normale Superiore di Pisa, Classe di Scienze, in press.

[7] G. ALBERTI, S. BIANCHINI & G. CRIPPA: *On the L^p differentiability of certain classes of functions.* Revista Matemática Iberoamericana, in press.

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

[9] L. AMBROSIO, N. FUSCO & D. PALLARA: *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, 2000.

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

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

[12] L. AMBROSIO: *Transport equation and Cauchy problem for non-smooth vector fields.* Lecture Notes in Mathematics "Calculus of Variations and Non-Linear Partial Differential Equations" (CIME Series, Cetraro, 2005) **1927**, B. Dacorogna, P. Marcellini eds., 2–41, 2008.

[13] L. AMBROSIO & G. CRIPPA: *Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields.* Lecture Notes of the UMI, **5** (2008), 3–54.

[14] 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.

[15] L. AMBROSIO, G. CRIPPA & S. MANIGLIA: *Traces and fine properties of a BD class of vector fields and applications.* Ann. Sci. Toulouse, **XIV** (4) (2005), pp. 527–561.

[16] 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, second edition in 2008.

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

[18] L. AMBROSIO & J. MALÝ: *Very weak notions of differentiability.* Proceedings of the Royal Society of Edinburgh, **137 A** (2007), 447–455.

[19] L. AMBROSIO, C. DE LELLIS & J. MALÝ: *On the chain rule for the divergence of BV like vector fields: applications, partial results, open problems.* In "Perspectives in Nonlinear Partial Differential Equations: in honour of Haim Brezis", Contemporary Mathematics, **446** (2007), 31–67, AMS.

[20] L. AMBROSIO, S. LISINI & G. SAVARÉ: *Stability of flows associated to gradient vector fields and convergence of iterated transport maps.* Manuscripta Mathematica, **121** (2006), 1–50.

[21] L. AMBROSIO & A. FIGALLI: *Geodesics in the space of measure-preserving maps and plans.* Arch. Rational Mech. Anal., **194** (2009), 421–462.

[22] L. AMBROSIO & A. FIGALLI: *On flows associated to Sobolev vector fields in Wiener spaces: an approach à la DiPerna-Lions.* Journal of Functional Analysis, **256** (2009), 179–214.

[23] L. AMBROSIO, M. COLOMBO, G. DE PHILIPPIS, & A. FIGALLI: *Existence of Eulerian solutions to the semigeostrophic equations in physical space: the two-dimensional periodic case.* Comm. Partial Differential Equations, **37** (2012), 2209–2227.

[24] L. Ambrosio, M. Colombo, G. De Philippis & A. Figalli: *A global existence result for the semigeostrophic equations in three dimensional convex domains.* Preprint, 2012, to appear on Discrete and Continuous Dynamical Systems.

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

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

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

[28] 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.

[29] P. Bernard & B. Buffoni: *Optimal mass transportation and Mather theory.* Journal of the European Mathematical Society, **1** (2007), 85–121.

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

[31] 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.

[32] A. Bohun, F. Bouchut & G. Crippa: Work in progress, 2013.

[33] A. Bohun, F. Bouchut & G. Crippa: Work in progress, 2013.

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

[35] F. Bouchut & G. Crippa: *Equations de transport à coefficient dont le gradient est donné par une intégrale singulière. (French) [Transport equations with a coefficient whose gradient is given by a singular integral].* Séminaire: Équations aux Dérivées Partielles. 2007–2008, Exp. No. I, 15 pp., Sémin. Équ. Dériv. Partielles, École Polytech., Palaiseau, 2009.

[36] F. Bouchut & G. Crippa: *Lagrangian flows for vector fields with gradient given by a singular integral.* J. Hyperbolic Differ. Equ., **10** (2013), 235–282.

[37] 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.

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

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

[40] 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.

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

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

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

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

[45] A. Bressan: *A lemma and a conjecture on the cost of rearrangements.* Rend. Sem. Mat. Univ. Padova, **110** (2003), 97–102.

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

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

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

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

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

[51] 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.

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

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

[54] N. Champagnat & P.-E. Jabin: *Well posedness in any dimension for Hamiltonian flows with non BV force terms.* Comm. Partial Differential Equations **35** (2010), 786–816.

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

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

[57] F. COLOMBINI, T. LUO & J. RAUCH: *Uniqueness and nonuniqueness for nonsmooth divergence-free transport.* Seminaire: Équations aux Dérivées Partielles, 20022003, Exp. No. XXII, 21 pp., Sémin. Équ. Dériv. Partielles, École Polytech., Palaiseau, 2003.

[58] G. CRIPPA: *The flow associated to weakly differentiable vector fields.* Theses of Scuola Normale Superiore di Pisa (New Series), **12**. Edizioni della Normale, Pisa, 2009 - Distributed by Birkhäuser.

[59] G. CRIPPA: *Ordinary Differential Equations and Singular Integrals.* HYP2012 Proceedings, in press, 2012.

[60] G. CRIPPA & C. DE LELLIS: *Oscillatory solutions to transport equations.* Indiana Univ. Math. J., **55** (2006), 1–13.

[61] G. CRIPPA & C. DE LELLIS: *Estimates for transport equations and regularity of the DiPerna-Lions flow.* J. Reine Angew. Math., **616** (2008), 15–46.

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

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

[64] M. CULLEN & M. FELDMAN: *Lagrangian solutions of semigeostrophic equations in physical space.* J. Math. Anal., **37** (2006), 1371–1395.

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

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

[67] L. DE PASCALE, M. S. GELLI & L. GRANIERI: *Minimal measures, one-dimensional currents and the Monge-Kantorovich problem.* Calculus of Variations and Partial Differential Equations, **27** (2006), 1–23.

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

[69] G. DE PHILIPPIS & A. FIGALLI: $W^{2,1}$ *regularity for solutions to the Monge-Ampére equation.* Invent. Math., **192** (2013), 55–69.

[70] G. DE PHILIPPIS, A. FIGALLI & O. SAVIN: *A note on $W^{2,1+\epsilon}$ interior regularity for solutions to the Monge-Ampére equation.* Math. Ann., **357** (2013), 11-22.

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

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

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

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

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

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

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

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

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

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

[81] P.-E. Jabin: *Differential equations with singular fields.* J. Math. Pures Appl. **94** (2010), 597–621.

[82] 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.

[83] 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.

[84] N. Lerner: *Transport equations with partially BV velocities.* Annali Scuola Normale Superiore, **3** (2004), 681–703.

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

[86] 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.

[87] 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.

[88] J. Lott & C. Villani: *Weak curvature conditions and Poincaré inequalities.* J. Funct. Anal., **245** (2007), 311–333.

[89] S. Maniglia: *Probabilistic representation and uniqueness results for measure-valued solutions of transport equations.* J. Math. Pures Appl., **87** (2007), 601–626.

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

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

[92] 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.

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

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

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

[96] T. Schmidt: $W^{1,2+\epsilon}$ *estimates for solutions to the Monge-Ampére equation.* Adv. Math., **240** (2013), 672–689.

[97] 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.

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

[99] 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.

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

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

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

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

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

[105] C. Villani: *Optimal transport: old and new.* Vol. 338 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.

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