Transport Equations and Flows with One-Sided Lipschitz Velocity Fields

PIERRE-LOUIS LIONS & BENJAMIN SEEGER

Communicated by C. LE BRIS

Abstract

We study first- and second-order linear transport equations, as well as flows for ordinary and stochastic differential equations, with irregular velocity fields satisfying a one-sided Lipschitz condition. Depending on the time direction, the flows are either compressive or expansive. In the compressive regime, we characterize the stable continuous distributional solutions of both the first and second-order nonconservative transport equations as the unique viscosity solution, and we also provide new observations and characterizations for the dual, conservative equations. Our results in the expansive regime complement the theory of Bouchut et al. (Ann Sc Norm Super Pisa Cl Sci (5) 4:1–25, 2005), and we develop a complete theory for both the conservative and nonconservative equations in Lebesgue spaces, as well as proving the existence, uniqueness, and stability of the regular Lagrangian flow for the associated ordinary differential equation. We also provide analogous results in this context for second order equations with degenerate noise coefficients that are constant in the spatial variable, as well as for the related stochastic differential equation flows.

Mathematics Subject Classification: 35F10 · 34A36 · 35A02 · 60H10 · 60H30

Contents

Pierre-Louis Lions and Benjamin Seeger contributed equally to the paper.

1. Introduction

For a fixed, finite time horizon $T > 0$ and a velocity field $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, we study the linear transport equation

$$\partial_t u + b(t, x) \cdot \nabla u = 0 \quad \text{in } [0, T] \times \mathbb{R}^d, \tag{1.1}$$

along with the dual, continuity equation

$$\partial_t f + \operatorname{div}(b(t, x) f) = 0 \quad \text{in } [0, T] \times \mathbb{R}^d, \tag{1.2}$$

and the associated ordinary differential equation (ODE) flow

$$\partial_t \phi_{t,s}(x) = b(t, \phi_{t,s}(x)), \quad (s, t, x) \in [0, T] \times [0, T] \times \mathbb{R}^d, \quad \phi_{s,s} = \operatorname{Id}. \tag{1.3}$$

The goal of the paper is to analyze the three problems, and the relations between them, for vector fields b satisfying the one-sided Lipschitz condition

$$\begin{cases} (b(t, x) - b(t, y)) \cdot (x - y) \\ \qquad \geq -C(t)|x - y|^2 \quad \text{for a.e. } (t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \\ \text{for some nonnegative } C \in L^1([0, T]). \end{cases} \tag{1.4}$$

When b is Lipschitz continuous in the space variable, the ODE flow (1.3) admits a unique global solution, and, through the method of characteristics, (1.1) and (1.2) are uniquely solved for any given smooth initial or terminal data. Moreover, the flow is a diffeomorphism, and therefore the solution operators for either the initial value problem (IVP) or terminal value problem (TVP) for (1.1) and (1.2) are continuous on L_{loc}^p for any $p \in [1, \infty]$.

Under the assumption (1.4), the time direction plays a nontrivial role, and there is a fundamental difference between the solvability of the flow (1.3) forward versus backward in time. Indeed, b need not even be continuous, and (1.4) is equivalent to

$$\frac{\nabla b(t, \cdot) + \nabla b(t, \cdot)^T}{2} \geq -C(t) \, \text{Id} \quad \text{in the sense of distributions.}$$

In particular, the distribution div$b$ is a signed measure that is bounded from below, but not in general absolutely continuous with respect to Lebesgue measure. Thus, when $t < s$, the flow (1.3) is expected to concentrate at sets of Lebesgue measure zero, while the formation of vacuum is witnessed for $t > s$.

A general study of transport equations and ODEs with irregular velocity fields, motivated by nonlinear problems in fluid dynamics, was initiated by DiPerna and the first author [40], who introduced the notion of renormalized solutions to prove the well-posedness for (1.1) and (1.2) and the almost-everywhere solvability of the flow (1.3) for b with Sobolev regularity. The DiPerna–Lions theory was extended to equations where only $\text{Sym}(\nabla b) \in L^1$ [28], to Vlasov equations with BV_{loc} velocity fields [20], and to two-dimensional problems with a Hamiltonian structure [2–4,21,45]. Using deep results from geometric measure theory, the renormalization property was extended to the very general case where $b \in BV_{\text{loc}}$ and $\text{div} \, b \in L^1$ by Ambrosio [5], who also provided a new, measure-theoretic viewpoint on the relationship between uniqueness of nonnegative solutions of (1.2) and the unique solvability of the flow (1.3) through the idea of superposition. Further developments include equations with velocity fields having a particular structure allowing for less regularity [31,50] and velocity fields belonging to SBD (i.e. $\text{Sym}(\nabla b)$ is a signed measure with no singular Cantor-like part) [12]. Fine regularity properties of DiPerna-Lions flows were established in [13,36], and the study of so-called "nearly incompressible flows" [14] led to the resolution by Bianchini and Bonicatto [19] of Bressan's compactness conjecture [26,27]; see also [47] for related results. For many more details and references, we refer the reader to the surveys [6–8,10].

In the majority of these works, the divergence div b is assumed to be bounded, or at least absolutely continuous with respect to Lebesgue measure. This is not the case in general for velocity fields satisfying (1.4), and so the equations (1.1) and

(1.2) do not even have a sense as distributions, because the products $(\operatorname{div} b)u$ and bf are ill-defined for general $u \in L^1_{loc}$ or measures f. The DiPerna-Lions theory does not, therefore, cover this situation. Moreover, the choice of an appropriate function space of solutions is very sensitive to whether the equations are posed as initial or terminal value problems.

The problems (1.1)–(1.3) for velocity fields with a one-sided Lipschitz condition have been approached with a variety of methods [22,25,29,33,59–61], a primary motivation being the study of pressureless gases and scalar conservation laws, which, when posed as nonlinear transport equations, involve velocity fields whose divergence is not absolutely continuous [23,24,43,44]. Our main purpose is to complement these works, and in particular the theory of Bouchut, James, and Mancini [25], by providing complete characterizations of the stable solutions to all three problems in both the compressive and expansive regimes. We also provide some results on the corresponding parabolic equations with a degenerate, second-order term, as well as the SDE analogue of (1.3) for both the velocity field b and $-b$.

1.1. Main Results

We relegate a full description of the results, discussions, and examples to the body of the paper. Here, we briefly outline the different sections and the types of results proved within them, and we compare them to the existing literature.

1.1.1. The Compressive Regime   In Sect. 2, we record properties of the backward Filippov flow for (1.3), as well as for its Jacobian $J_{t,s}(x) := \det(\nabla \phi_{t,s}(x))$, which is well-defined in L^∞ for a.e. $t \le s$ and $x \in \mathbb{R}^d$. We employ measure-theoretic arguments to make sense of the right-inverse of the flow in an almost-everywhere sense, as a preliminary step to understanding the forward, regular Lagrangian flow, and prove several properties, the most important of which is its almost-everywhere continuity.

In Sect. 3, we turn to the study of the nonconservative equation[1]

$$\partial_t u - b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T, \tag{1.5}$$

for which the uniqueness of continuous distributional solutions fails in general. We introduce a new PDE characterization of the "good" (stable) solution of (1.5) as the unique viscosity solution, in the sense of Crandall, Ishii, and the first author [35]. This is done by proving a comparison principle for sub and supersolutions. The viscosity solution characterization coincides with the selection of "good" solutions

---

[1] For a consistent presentation throughout the paper, and in order to emphasize the dual relationship between the two equations, the transport equation (1.1) will always be posed as a terminal value problem, and the continuity equation (1.2) as an initial value problem. The compressive and expansive regimes will be distinguished by the choice of sign in front of the velocity field b.

by other authors in particular settings [22,25,33,59–61], allows for robust stability statements, and, moreover, generalizes to the setting of degenerate parabolic problems (see the discussion below).

The "usual" viscosity solution theory must be modified due to the lack of global continuity of b. In view of the evolution nature of the equations, the L^1-dependence in time does not present a problem, and the equations can be treated with the methods of [46,55,57,58]. To deal with the discontinuity of b in space, sub and supersolutions must be defined with appropriate semicontinuous envelopes of b in the space variable. The direction of the one-sided Lipschitz assumption (1.4) accounts for the beneficial inequalities in the proof of the comparison principle.

The nonuniqueness of distributional solutions is explored through examples of the form $b(x) = \operatorname{sgn} x |x|^\alpha$. We also introduce further conditions on the velocity field b and terminal data u_T that ensure uniqueness of arbitrary continuous distributional solutions. In particular, the interplay between the regularity of b and u_T plays an important role: if $b \in C^\alpha$ and $u_T \in C^\beta$, then distributional solutions are unique if $\alpha + \beta > 1$, while uniqueness may fail in general if $\alpha + \beta \le 1$, as can be seen from our counterexamples.

The latter half of Sect. 3 deals with the study of the dual problem to (1.5), namely,

$$\partial_t f - \operatorname{div}(b(t,x)f) = 0 \quad \text{in } (0,T) \times \mathbb{R}^d, \quad f(0,\cdot) = f_0. \qquad (1.6)$$

Even if $f_0 \in L^1_{\mathrm{loc}}$, the concentrative nature of the flow causes the measure $f(t,\cdot)$ to develop a singular part, and therefore we are led to seek measure-valued solutions. This prevents the duality solution of (1.6) from being understood in the distributional sense, due to the lack of continuity of b. Nevertheless, we prove that, if b is continuous, or if it happens that $f(t,\cdot)$ is absolutely continuous with respect to Lebesgue measure on the time interval $[0,T]$, then the notions of duality and distributional solutions are equivalent.

An important feature of the continuity equation (1.6) is the failure of renormalization; that is, if f is a duality solution, the measure $|f|$ may fail to be a distributional solution, and may even violate conservation of mass. We once again study examples of the form $b(x) = \operatorname{sgn} x |x|^\alpha$, $0 \le \alpha < 1$. Note that, for this example, when $\alpha > 1$, b has the Sobolev regularity $b \in W^{1,p}$ for $p < 1/\alpha$, and so our counterexample is constructed to ensure that the duality solution f satisfies $f \in L^q$ only for q outside the range for which the DiPerna-Lions commutator lemma holds. This contrast with the DiPerna-Lions theory is a direct consequence of the compressive nature of the backward flow, which can lead to cancellation of the positive and negative parts of f. A related phenomenon is the nonuniqueness of distributional solutions of the continuity equation (1.2) with the reverse sign (see below).

1.1.2. The Expansive Regime   In Sect. 4, we reverse the sign on the velocity field, and study the corresponding problems

$$\partial_t u + b(t,x) \cdot \nabla u = 0 \quad \text{in } (0,T) \times \mathbb{R}^d, \quad u(T,\cdot) = u_T \qquad (1.7)$$

and

$$\partial_t f + \operatorname{div}(b(t,x)f) = 0 \quad \text{in } (0,T) \times \mathbb{R}^d, \quad f(0,\cdot) = f_0. \qquad (1.8)$$

In view of the lower bound on the divergence of b, we are motivated to seek an L^p-based theory for both equations, based on a priori estimates, or equivalently, on the fact that the characteristic flow (the forward ODE (1.3)) does not concentrate on sets of measure zero.

The initial value problem for the continuity equation (1.8) was studied in [22, 25], where a large part of the analysis is based on the fact that locally integrable distributional solutions are *not* unique in general.[2] The same setting is studied in [29], where the existence and uniqueness of the forward Filippov flow for (1.3) is established for a.e. $x \in \mathbb{R}^d$.

In the first part of Sect. 4, we identify a unique "good" distributional solution, and prove that the resulting solution operator is continuous on L^p_{loc} for all $p \in [1, \infty]$, and stable with respect to regularizations. This coincides with the notion of reversible solution in [22,25].

We then obtain strong stability results for the Bouchut–James–Mancini duality solutions of the nonconservative problem (1.7) in all L^p-spaces, which allow us to prove the renormalization property. Moreover, we introduce a PDE characterization of this duality solution in terms of regularization by ess inf- and ess sup-convolution. An important ingredient in establishing this characterization is the propagation of almost-everywhere continuity, which, in turn, follows from the renormalization property and the almost-everywhere continuity of the forward flow proved in Sect. 2.

As a consequence of this new characterization, we give a PDE-based proof of the fact that *nonnegative* distributional L^p-solutions of (1.8) are unique, which was established in [29] using the superposition principle. This result, along with the renormalization property for (1.7), allows us to establish the existence, uniqueness, and stability of the forward regular Lagrangian flow for the ODE (1.3) identified in [29]. As a byproduct, this also provides a full characterization of the Bouchut-James-Mancini notion of "good" (reversible) solution as the pushforward of f_0 by the forward flow. Moreover, a distributional solution f is a reversible solution if and only if $|f|$ is also a distributional solution (cf. [25, Proposition 3.12], which operates under the criterion that f be a so-called "Jacobian" solution).

1.1.3. SDEs and Second Order Equations

This paper also contains various results regarding second order versions of (1.1) and (1.2), as well as stochastic differential equation (SDE) flows. SDEs and degenerate second-order Fokker-Planck equations have been studied from many perspectives, using both the DiPerna-Lions theory and adaptations of the superposition principle, by many authors, including Le Bris and Lions [51], Figalli [41], Trevisan [64], and Champagnat and Jabin [32]; see also the book [52]. Just as in the first-order setting, the fact that the measure div b may contain a singular part prevents the application of these theories to the present situation.

---

[2] Note that in [25], the velocity field a takes the role of $-b$ in (1.4), while the time-directions of (1.7) and (1.8) are switched, so their setting corresponds to the expansive regime discussed here.

In the compressive regime, we extend the viscosity solution theory of Sect. 3 to the second order equation

$$-\partial_t u + b(t, x) \cdot \nabla u - \text{tr}[a(t, x)\nabla^2 u] = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T, \quad (1.9)$$

where b satisfies the one-sided Lipschitz condition (1.4) and a is a regular, but possibly degenerate, symmetric matrix. This equation, as well as the dual problem

$$\partial_t f - \text{div}(b(t, x)f) - \nabla^2 \cdot (a(t, x)f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad f(0, \cdot) = f_0, \quad (1.10)$$

can be related to the SDE

$$d_t \Phi_{t,s}(x) = -b(t, \Phi_{t,s}(x))dt + \sigma(t, \Phi_{t,s}(x))dW_t, \quad t > s, \quad \Phi_{s,s}(x) = x,$$
$$(1.11)$$

which is the SDE analogue of the backward flow for (1.3). Here W is a given Brownian motion and $a = \frac{1}{2}\sigma\sigma^T$. We establish the existence and uniqueness, for every $x \in \mathbb{R}^d$, of a strong solution in the Filippov sense, and we show that, with probability one, $\Phi_{t,s}$ is Hölder continuous for any exponent less than 1.

The situation is more complicated in the expansive regime, namely, for the equations

$$-\partial_t u - b(t, x) \cdot \nabla u - \text{tr}[a(t, x)\nabla^2 u] = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T \quad (1.12)$$

and

$$\partial_t f + \text{div}(b(t, x)f) - \nabla^2 \cdot (a(t, x)f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad f(0, \cdot) = f_0. \quad (1.13)$$

In the first-order setting, the characterization of the "good" distributional solution of the continuity equation (1.8) relies on the Lipschitz continuity of the backward ODE flow. Adapting similar methods for the second order equation (1.13) involves establishing Lipschitz continuity of a stochastic flow like (1.11) with certain time-reversed coefficients (see (4.30) below). While it is well-known that flows of the form (1.11) are Hölder continuous for any exponent less than 1, even in more general contexts (see [62]), it is an open question as to whether it is Lipschitz with probability one. We relegate a general study of (1.12) and (1.13), and of the stochastic regular Lagrangian flow for

$$d_t \Phi_{t,s}(x) = b(t, \Phi_{t,s}(x))dt + \sigma(t, \Phi_{t,s}(x))dW_t, \quad t > s, \quad \Phi_{s,s}(x) = x, \quad (1.14)$$

to future work. The exception is when σ is constant in the $\mathbb{R}^d$-variable.[3] In this case, we prove that a suitable stochastic flow of the form (1.11) can be inverted, leading, as in the deterministic case, to the existence and uniqueness of a strong solution to (1.14) for a.e. $x \in \mathbb{R}^d$, and a corresponding solution theory for the PDEs (1.12) and (1.13).

---

[3] Another case of interest is when the diffusion matrix a is nondegenerate in which case very general results can be obtained even for locally bounded b; see [41].

1.2. Applications and Further Study

While interesting in their own right, linear transport equations and ODEs with nonregular velocity fields arise naturally in several equations in fluid dynamics, in which the velocity fields depend nonlinearly on various other physical quantities that are coupled with the transported quantity. Since these equations must be posed a priori in a weak sense, this leads to velocity fields with limited regularity. The DiPerna-Lions and Ambrosio theories have been successfully applied to a number of such situations; see [9,11,15,16,37,48,53,67]. As mentioned above, the one-dimensional Bouchut-James theory of reversible solutions for transport equations with semi-Lipschitz velocity fields has been successfully applied in applications to conservation laws and pressureless gasses; see [23,24,43,44].

Nonlinear transport equations also arise in certain models for large population dynamics, specifically mean field games (MFG). In [49], the first author and Lasry introduced a forward-backward system of PDEs modeling a large population of agents in a state of Nash equilibrium. The evolution of the density f of players is described by a continuity equation (1.8) (or Fokker-Planck equation (1.13)), where the velocity field b is given by

$$b(t, x) = -\nabla_p H(t, x, \nabla u(t, x)). \tag{1.15}$$

Here, H is a convex Hamiltonian, and u is the solution of the terminal value problem

$$-\partial_t u - \mathrm{tr}[a(t, x)\nabla^2 u] + H(t, x, \nabla u(t, x)) = F[f(t, \cdot)] \quad \text{in } (0, T) \times \mathbb{R}^d,$$
$$u(T, \cdot) = G[f(t, \cdot)], \tag{1.16}$$

which is a Hamilton–Jacobi–Bellman equation encoding the optimization problem for a typical agent, and whose influence by the population of agents is described by the coupling functions F and G. The velocity field (1.15) is the consensus optimal feedback policy of the population of agents at a Nash equilibrium.

When a is degenerate, or even zero, the function u has limited regularity, and is no better than semiconcave in the spatial variable in general. Therefore, even if H is smooth, the velocity field (1.15) may satisfy at most

$$b \in BV_{\mathrm{loc}} \quad \text{and} \quad (\mathrm{div}\, b)_- \in L^\infty. \tag{1.17}$$

This falls just outside the DiPerna-Lions-Ambrosio regime, since the measure $(\mathrm{div}\, b)_+$ may still fail to be absolutely continuous in general. In fact, the well-posedness of a suitable notion of solution for the transport and ODE problems under the general assumptions (1.17) remains an open problem.

Many simple but useful MFG models involve a linear-quadratic Hamiltonian of the form

$$H(t, x, p) = A(t, x)|p|^2 + B(t, x) \cdot p + C(t, x)$$

for smooth, real-valued A, B, C with $A > 0$. In this case, it is easy to see that (1.15) satisfies the half-Lipschitz condition (1.4). This situation was studied by Cardaliaguet and Souganidis [29] for first-order, stochastic mean field games systems with common noise. In particular, it is proved there that the uniqueness of

probability density solutions of (1.7) gives rise, through the superposition princi-
ple, to the uniqueness of optimal trajectories for the probabilistic formulation of
the MFG problem, and, moreover, the solution of the stochastic forward-backward
system can be used to construct approximate Nash equilibria for the N-player game.
Our analysis for the Fokker-Planck equation (1.13) may therefore be expected to
yield similar results for stochastic MFG systems with common noise and degen-
erate, spatially-homogenous, idiosyncratic noise, a special case of the equations
considered by Cardaliaguet, Souganidis, and the second author in [30].

The second application of nonlinear transport equations in mean field games
is involved with the master equation for a MFG with a finite state space. These
equations generally take the form

$$\partial_t u + b(t, x, u) \cdot \nabla u = c(t, x, u) \quad \text{in } (0, T) \times \mathbb{R}^d, \tag{1.18}$$

where u, b, and c all take values in $\mathbb{R}^d$; coordinate-by-coordinate, (1.18) is written
as

$$\partial_t u^i + b^j(t, x, u)\partial_{x_j} u^i = c^i(t, x, u), \quad i = 1, 2, \ldots, d.$$

Therefore, (1.18) is a nonconservative hyperbolic system, whose general well-
posedness is a difficult question in general; note that, when $d = 1$, (1.18) becomes
a scalar conservation law.

We do not discuss (1.18) here, but, in the paper [56], we study a particular
regime of equations taking the form (1.18), using a new theory for linear transport
equations with velocity fields b that are increasing *coordinate by coordinate*, that
is, $\partial_{x_j} b^i \geq 0$ for $i \neq j$.

The extension to infinite dimensions, of both the linear problems (1.1)–(1.2),
as well as the nonlinear equation (1.18), remains an interesting question, with
numerous applications, including the study of mean field game master equations
on the Hilbert space of square-integrable random variables. We aim to study these
situations in future work.

1.3. Notation

Given a function space $X(\mathbb{R}^d)$, or $X(\Omega)$ for an appropriate subdomain of $\mathbb{R}^d$,
X_{loc} denotes the space of functions (or distributions) f such that $\phi f \in X$ for all
$\phi \in C_c^\infty(\mathbb{R}^d)$. If X is a normed space, the same is not necessarily true for X_{loc}, but
it inherits the topology of local X-convergence. For example, $\lim_{n\to\infty} f_n = f$ in
$L^p_{\text{loc}}(\mathbb{R}^d)$ means that $\lim_{n\to\infty} \|f_n - f\|_{L^p(B_R)} = 0$ for all $R > 0$. We denote by
$L^p_+([0, T])$ the subset of $L^p([0, T])$ consisting of nonnegative functions.

Unless otherwise specified, Banach or Fréchet spaces of functions are endowed
with the strong topology. For a function space X, the subscripts X_{w} and $X_{\text{w-}\star}$
indicate the weak (resp. weak-$\star$) topology.

For $1 \leq p < \infty$, $\mathcal{P}_p$ is the space of Borel probability measures μ, with
$\int |x|^p \mu(dx) < \infty$, which becomes a complete metric space for the p-Wasserstein
distance $\mathcal{W}_p$ defined for $\mu, \nu \in \mathcal{P}_p$ by

$$\mathcal{W}_p(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma(x, y),$$

where Γ is the set of couplings of μ and ν, that is, measures γ on the product space $\mathbb{R}^d \times \mathbb{R}^d$ such that $\gamma(A \times \mathbb{R}^d) = \mu(A)$ and $\gamma(\mathbb{R}^d \times A) = \nu(A)$ for all Borel measurable $A \subset \mathbb{R}^d$.

The transpose of a matrix σ is denoted by σ^T, and, if σ is a square matrix, its symmetric part is denoted by $\mathrm{Sym}(\sigma) := \frac{1}{2}(\sigma + \sigma^T)$. The symbol Id stands for either the identity map or the identity matrix, the precise meaning being clear from context.

2. The ODE Flow

This section is focused on the solvability and properties of the flow associated to a velocity field b satisfying that

$$\begin{cases} \text{for some } C_0, C_1 \in L^1_+([0, T]) \text{ and for all } t \in [0, T] \text{ and } x, y \in \mathbb{R}^d, \\ |b(t, x)| \leq C_0(t)(1 + |x|) \quad \text{and} \\ (b(t, x) - b(t, y)) \cdot (x - y) \geq -C_1(t)|x - y|^2. \end{cases} \quad \text{[4]} \quad (2.1)$$

Because $b(t, \cdot)$ is not necessarily continuous, the ODE must be interpreted in the Filippov sense [42], that is, abusing notation, we denote by $b(t, x)$ the convex hull of all limit points of $b(t, y)$ as $y \to x$. For $s \in [0, T]$, we seek absolutely continuous solutions $t \mapsto \phi_{t,s}(x)$ of the problem

$$\begin{cases} \partial_t \phi_{t,s}(x) \in b(t, \phi_{t,s}(x)), \ t \in [0, T], \\ \phi_{s,s}(x) = x. \end{cases} \quad (2.2)$$

Remark 2.1. If $\dot{X}(t) \in b(t, X(t))$,

$$\tilde{X}(t) := \exp\left(\int_0^t C_1(s)ds\right) X(t) \quad \text{and}$$

$$\tilde{b}(t, x) := \exp\left(\int_0^t C_1(s)ds\right) b\left(t, \exp\left(-\int_0^t C_1(s)ds\right) x\right),$$

so that $\dot{\tilde{X}} \in \tilde{b}(t, \tilde{X}(t))$, then $\tilde{b}$ satisfies (2.1) with $C_1 \equiv 0$ and a possibly different C_0. In other words, with a change of variables, one may always assume b is monotone without loss of generality.

We will use the following characterization and properties of half-Lipschitz maps; see [25, Lemma 2.2 and Remark 2.4].

---

[4] The linear growth assumption is a standard way to ensure that the a priori estimates for solutions do not blow up. Otherwise, the results of the paper would need a corresponding local theory, as for example in [17].

Lemma 2.1. *A vector field $B : \mathbb{R}^d \to \mathbb{R}^d$ satisfies*

$$(B(x) - B(y)) \cdot (x - y) \geq -C|x - y|^2 \quad \text{for some } C \geq 0 \text{ and all } x, y \in \mathbb{R}^d$$

if and only if $\mathrm{Sym}(\nabla B) \geq -C \, \mathrm{Id}$ *in the sense of distributions. We then also have* $B \in BV_{\mathrm{loc}}(\mathbb{R}^d)$, *and*

$$\mathrm{Sym}(\nabla B) - (\mathrm{tr}\, \nabla B)\, \mathrm{Id} \leq (d-1)C \, \mathrm{Id} \,.$$

The fact that B belongs not only to the space $BD_{\mathrm{loc}}(\mathbb{R}^d)$ of bounded deformations (the space of vector fields $B : \mathbb{R}^d \to \mathbb{R}^d$ such that the symmetric part of the distribution ∇B is a locally bounded Radon measure [63]), but more particularly to $BV_{\mathrm{loc}}(\mathbb{R}^d)$, is a consequence of the analysis in [1], where we refer the reader for many more fine geometric properties of (semi-)monotone functions.

We fix a family of regularizations such that

$$\begin{cases} (b^\varepsilon)_{\varepsilon > 0} \subset L^1([0, T], C^{0,1}(\mathbb{R}^d)), \quad \lim_{\varepsilon \to 0} b^\varepsilon = b \text{ a.e. in } [0, T] \times \mathbb{R}^d, \text{ and} \\ b^\varepsilon \text{ satisfies (2.1) uniformly in } \varepsilon > 0. \end{cases} \quad (2.3)$$

For example, we may take $b^\varepsilon(t, \cdot) = b(t, \cdot) * \rho_\varepsilon$ for $\rho_\varepsilon = \varepsilon^{-d} \rho(\cdot/\varepsilon)$, with $\rho \in C_+^\infty(\mathbb{R}^d)$, $\mathrm{supp}\,\rho \in B_1$, and $\int \rho = 1$.

2.1. The Backward Flow

We begin the analysis with the backward flow, that is, (2.2) for $t < s$. This is the time-direction for which the one-sided Lipschitz condition (2.1) yields a unique, Lipschitz flow. We record its properties here and refer to [33,42,59,60] for the proofs; see also the work of Dafermos [38] for the connection to generalized characteristics of conservation laws.

Lemma 2.2. *For every $(s, x) \in [0, T] \times \mathbb{R}^d$, there exists a unique solution $\phi_{t,s}(x)$ of (2.2) defined for $(t, x) \in [0, s] \times \mathbb{R}^d$, satisfying the Lipschitz bound*

$$|\phi_{t,s}(x) - \phi_{t,s}(y)| \leq \exp\left(\int_t^s C_1(r)dr\right)|x - y| \quad \text{for all } 0$$

$$\leq t \leq s \leq T \text{ and } x, y \in \mathbb{R}^d. \quad (2.4)$$

Moreover, there exists a constant $C > 0$ depending only on T and C_0 from (2.1) such that

$$|\phi_{t,s}(x)| \leq C(|x| + 1) \quad \text{for all } 0 \leq t \leq s \leq T \text{ and } x \in \mathbb{R}^d, \quad (2.5)$$

and

$$\begin{cases} |\phi_{t_1,s}(x) - \phi_{t_2,s,x}| \leq C(1 + |x|)|t_1 - t_2| \quad \text{and} \\ |\phi_{t,s_1}(x) - \phi_{t,s_2}(x)| \leq C(1 + |x|)|s_1 - s_2| \\ \text{for all } t_1, t_2 \in [0, s], \ s_2, s_2 \in [t, T], \text{ and } x \in \mathbb{R}^d. \end{cases} \quad (2.6)$$

For all $0 \leq r \leq s \leq t \leq T$, $\phi_{r,s} \circ \phi_{s,t} = \phi_{r,t}$. If $(b^\varepsilon)_{\varepsilon > 0}$ are regularizations satisfying (2.3), then the corresponding backward flows ϕ^ε converge locally uniformly as $\varepsilon \to 0$ to ϕ.

Remark 2.2. The a priori local boundedness and time-regularity estimates (2.5) and (2.6), depending only on C_0 and not C_1, do not require the half-Lipschitz assumption on $b(t, \cdot)$, and are therefore satisfied for any limiting solutions of the ODE when b satisfies the first condition in (2.1). On the other hand, the half-Lipschitz assumption is crucial for the Lipschitz continuity of the flow (2.4), as well as the uniqueness of the solution.

Remark 2.3. Consider the backward flow in $\mathbb{R}$ corresponding to $b(t, x) = b(x) = \operatorname{sgn} x$, which is given, for $x \in \mathbb{R}$ and $s < t$, by

$$\phi_{s,t}(x) = \begin{cases} x + (t - s) & \text{if } x < -t - s, \\ 0 & \text{if } |x| \leq t - s, \text{ and} \\ x - (t - s) & \text{if } x > t - s. \end{cases} \qquad (2.7)$$

This demonstrates that, in general, the trajectories of the backward flow may concentrate on sets of measures 0, in particular, where b has jump discontinuities.

We will often consider the examples $b(x) = \operatorname{sgn} x$ in subsequent parts of the paper in order to illustrate certain general phenomena and to present counterexamples. Note that, by Remark 2.1, one can consider similar examples for arbitrary $C_1 \in L_+^1([0, 1])$.

2.2. The Jacobian for the Backward Flow

In view of the Lipschitz regularity (2.4), $\nabla_x \phi_{t,s} \in L^\infty$ for $t \leq s$, and so we can define the Jacobian

$$J_{t,s}(x) := \det(\nabla_x \phi_{t,s}(x)) \quad \text{for } 0 \leq t \leq s \leq T \quad \text{and a.e. } x \in \mathbb{R}^d. \qquad (2.8)$$

Lemma 2.3. *Let J be defined as in (2.8). Then $J \geq 0$,*

$$\begin{cases} J_{\cdot,s} \in L^\infty([0, s] \times \mathbb{R}^d) \cap C([0, s], L^1_{\text{loc}}(\mathbb{R}^d)) \quad \forall s \in [0, T] \quad and \\ J_{t,\cdot} \in L^\infty([t, T] \times \mathbb{R}^d) \cap C([t, T], L^1_{\text{loc}}(\mathbb{R}^d)) \quad \forall t \in [0, T], \end{cases} \qquad (2.9)$$

$$\|J_{t,s}\|_{L^\infty} \leq \exp\left(d \int_t^s C_1(r)dr\right) \quad for all \ 0 \leq t \leq s \leq T, \qquad (2.10)$$

and, for all $R > 0$, there exists a modulus of continuity ω_R, which depends on b only through the constants C_0 and C_1 in (2.1), such that

$$\begin{cases} \|J_{t_1,s} - J_{t_2,s}\|_{L^1(B_R)} \leq \omega_R(|t_1 - t_2|) \quad for all \ t_1, t_2 \in [0, s] \quad and \\ \|J_{t,s_1} - J_{t,s_2}\|_{L^1(B_R)} \leq \omega_R(|s_1 - s_2|) \quad for all \ s_1, s_2 \in [t, T]. \end{cases} \qquad (2.11)$$

If $(b^\varepsilon)_{\varepsilon>0}$ are as in (2.3), $(\phi^\varepsilon)_{\varepsilon>0}$ are the corresponding solutions of (2.2), and, for $\varepsilon > 0$, $J^\varepsilon = \det(\nabla_x \phi^\varepsilon)$, then

$$\lim_{\varepsilon \to 0} J^\varepsilon_{\cdot,s} = J_{\cdot,s} \quad weak\text{-}\star \ in \ L^\infty([0, s] \times \mathbb{R}^d) \quad and$$

$$\lim_{\varepsilon \to 0} J^\varepsilon_{t,\cdot} = J_{t,\cdot} \quad weak\text{-}\star \ in \ L^\infty([t, T] \times \mathbb{R}^d). \qquad (2.12)$$

Proof. It suffices to prove all statements about $J_{\cdot,s}$ on $[0, s]$. The arguments are exactly the same for the other halves using the fact that $s \mapsto \phi_{t,s}$ is the forward flow corresponding to the velocity $-b$.

The convergence (2.12) goes through by compensated compactness arguments for determinants; see the Appendix of [25]. The nonnegativity of J now follows, because $J^\varepsilon \geq 0$ for all ε.

For fixed $\varepsilon > 0$ and $(s, x) \in [0, T] \times \mathbb{R}^d$, we have

$$\partial_t J^\varepsilon_{t,s}(x) = \operatorname{div}_x b^\varepsilon(t, \phi^\varepsilon_{t,s}(x)) J^\varepsilon_{t,s}(x) \quad \text{for } t \in [0, s].$$

Then (2.3) implies $\partial_t J^\varepsilon_{t,s}(x) \geq -dC_1(t) J^\varepsilon_{t,s}(x)$, and so

$$\frac{\partial}{\partial t}\left(J^\varepsilon_{t,s}(x) e^{-d \int_t^s C_1(r) dr} \right) \geq 0.$$

In particular, for $t_1 < t_2 \leq s$ and $R > 0$,

$$\int_{B_R} |J^\varepsilon_{t_2,s} - J^\varepsilon_{t_1,s}| \leq e^{\int_{t_1}^{t_2} C_1(r) dr} \int_{B_R} J^\varepsilon_{t_2,s} - \int_{B_R} J^\varepsilon_{t_1,s}$$
$$+ \left(e^{\int_{t_1}^{t_2} C_1(r) dr} - 1 \right) \int_{B_R} J^\varepsilon_{t_2,s}.$$

Identifying the modulus of continuity ω_R in the statement of the Lemma then reduces to proving the uniform-in-ε continuity of

$$[0, s] \ni t \mapsto \int_{B_R} J^\varepsilon_{t,s}(x) dx;$$

note that $\int_{B_R} J^\varepsilon_{s,s}(x) dx = |B_R|$, so this will also imply that $\int_{B_R} J^\varepsilon_{t,s}(x) dx$ is bounded uniformly in ε.

In view of the uniform-in-ε L^∞-boundedness of J^ε, it suffices to prove the uniform-in-ε continuity in t of $\int f(x) J^\varepsilon_{t,s}(x) dx$ for any $f \in C_c(\mathbb{R}^d)$. The change of variables formula gives

$$\int f(x) J^\varepsilon_{t,s}(x) dx = \int f(\phi^\varepsilon_{s,t}(x)) dx.$$

Note that $\partial_t \phi^\varepsilon_{s,t}(x) = -b^\varepsilon(t, \phi^\varepsilon_{s,t}(x))$, and the Lipschitz constant in t of $\phi^\varepsilon_{s,t}(x)$ depends only on an upper bound for $|x|$ and the constant C_0 in (2.1), and, therefore, is independent of ε. □

When $d = 1$, the L^∞-weak-$\star$ convergence of $J^\varepsilon = \partial_x \phi^\varepsilon$ to J can be strengthened via an Aubin-Lions type compactness result.

Proposition 2.1. *Assume $d = 1$, and let J^ε and J be as in Lemma 2.3. Then*

$$\lim_{\varepsilon \to 0} J^\varepsilon_{\cdot,s} = J_{\cdot,s} \quad \text{strongly in } L^1_{\text{loc}}([0, s] \times \mathbb{R})$$

and

$$\lim_{\varepsilon \to 0} J^\varepsilon_{t,\cdot} = J_{t,\cdot} \quad \text{strongly in } L^1_{\text{loc}}([t, T] \times \mathbb{R}).$$

Proof. Fix $t \in [0, T]$ and $R > 0$. Lemma 2.2 implies that there exists M independent of ε such that $|\phi_{t,s}^{\varepsilon}(x)| \leq M$ for all $s \in [t, T]$ and $x \in [-R, R]$. Upon redefining b outside of $[0, T] \times [-2R, 2R]$, we find that $\phi_{t,s}(x)$, and therefore $J_{t,s}(x)$, is unchanged, and therefore, in order to prove the L^1-convergence in $[t, T] \times [-R, R]$, we may assume without loss of generality that b is bounded uniformly. Applying the transformation $\tilde{\phi}_{t,s}(x) = \phi_{t,s}(x) - \int_t^s C(r) dr$ for an appropriate $C \in L_+^1([0, T])$ depending on C_0 from (2.1), we may also assume $b \geq 1$.

For $(s, x) \in [t, T] \times \mathbb{R}$, set $f^{\varepsilon}(s, x) = J_{t,s}^{\varepsilon}(x)$. Then f^{ε} solves the continuity equation

$$\partial_s f^{\varepsilon} + \partial_x \left(b^{\varepsilon}(s, x) f^{\varepsilon} \right) = 0 \quad \text{in } [t, T] \times \mathbb{R} \quad \text{and} \quad f^{\varepsilon}(t, \cdot) = 1.$$

For a standard mollifier $\rho \in C_c^{\infty}([-1, 1])$, let $\rho_n = n\rho(\cdot/n)$ and $f^{\varepsilon,n} = \rho_n *_t f^{\varepsilon}$ be the mollification of f^{ε} only in the time variable. We then have

$$\partial_s f^{\varepsilon,n} + \partial_x \left[\rho_n *_t (b^{\varepsilon} f^{\varepsilon}) \right] = 0 \quad \text{in } \left[t + \frac{1}{n}, T \right] \times \mathbb{R}$$

and, for any $R > 0$,

$$\sup_{s \in [t+1/n, T]} \left\| \partial_x \left[\rho_n *_t (b^{\varepsilon} f^{\varepsilon}) \right] (s, \cdot) \right\|_{L^1([-R,R])}$$

$$\leq \sup_{s \in [t+1/n, T]} \left\| \partial_s f^{\varepsilon,n}(s, \cdot) \right\|_{L^1([-R,R])} \leq n \left\| \rho' \right\|_{L^1(\mathbb{R})} \omega_R \left(\frac{1}{n} \right),$$

where ω_R is as in (2.11). It follows that, for fixed $n \in \mathbb{N}$, $(\rho_n *_t (b^{\varepsilon} f^{\varepsilon}))_{\varepsilon > 0}$ is precompact in $L^1([t, T] \times [-R, R])$, and so, because

$$\lim_{n \to \infty} \rho_n *_t (b^{\varepsilon} f^{\varepsilon}) = bf$$

in $L^1([t, T] \times [-R, R])$, uniformly in ε, we conclude that $(b^{\varepsilon} f^{\varepsilon})_{\varepsilon > 0}$ is precompact in $L^1([t, T] \times [-R, R])$. This implies that, as $\varepsilon \to 0$, $b^{\varepsilon} f^{\varepsilon}$ converges strongly in $L^1([t, T] \times [-R, R])$ to bf.

Fix any subsequence $(\varepsilon_n)_{n \geq 0}$ approaching zero as $n \to \infty$. Then there exists a further subsequence such that $f^{\varepsilon_{n_k}} b^{\varepsilon_{n_k}} \xrightarrow{k \to \infty} fb$ almost everywhere, and therefore $f^{\varepsilon_{n_k}} \xrightarrow{k \to \infty} f$ a.e. in $[t, T] \times [-R, R]$ because $b \geq 1$ and

$$f^{\varepsilon}(s, x) - f(s, x) = \frac{b(s, x) f^{\varepsilon}(s, x) - b(s, x) f(s, x)}{b(s, x)}$$

$$= \frac{b^{\varepsilon}(s, x) f^{\varepsilon}(s, x) - b(s, x) f(s, x)}{b(s, x)}$$

$$+ \frac{(b(s, x) - b^{\varepsilon}(s, x)) f^{\varepsilon}(s, x)}{b(s, x)}.$$

The convergence of $f^{\varepsilon_{n_k}}$ to f in $L^1([t, T] \times [-R, R])$, and therefore the convergence of the full family $(f^{\varepsilon})_{\varepsilon > 0}$ to f, is a consequence of the Lebesgue dominated convergence theorem. $\square$

Remark 2.4. The one-dimensional structure is important in the proof of Proposition 2.1, in particular, in deducing from the equicontinuity of J^ε in time that $(b^\varepsilon J^\varepsilon)_{\varepsilon>0}$ belongs to a precompact subset of L^1. It is not immediately clear whether this argument can be extended to multiple dimensions.

2.3. The Forward Flow as the Right-Inverse of the Backward Flow

We next investigate the solvability of (2.2) forward in time. This is done by analyzing the Jacobian J from the previous subsection in order to invert the backward flow. Similar methods are used in [29], and, by including the Jacobian in the analysis, we obtain additionally the almost-everywhere continuity of the inverse.

We will revisit this topic in Sect. 4 when we analyze the forward flow, which will arise from the theory of renormalized solutions of the appropriate transport equation.

Proposition 2.2. *For $t \leq s$, there exists a set $A_{ts} \subset \mathbb{R}^d$ of full measure such that, for all $y \in A_{ts}$, $\phi_{t,s}^{-1}(\{y\})$ is a singleton, which we denote by $\{\phi_{s,t}(y)\}$. Moreover, there exists a version of the map $\phi_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ such that $\phi_{s,t}$ is continuous a.e.*

As an intermediate step, we first prove the following:

Lemma 2.4. *Assume $0 \leq t \leq s \leq T$ and $K \subset \mathbb{R}^d$ is nonempty, compact, and connected. Then $\phi_{t,s}^{-1}(K)$ is nonempty, compact, and connected.*

Proof. For $r > 0$, define $K_r := \bigcup_{y \in K} B_r(y)$. Fix a sequence $(b^n)_{n \in \mathbb{N}}$ satisfying (2.3)[5], and let $\phi_{t,s}^n$ denote the corresponding backward flow from the previous subsections.

We first show that

$$\phi_{t,s}^{-1}(K) = \bigcap_{r>0} \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} (\phi_{t,s}^k)^{-1}(K_r). \qquad (2.13)$$

Suppose $x \in \phi_{t,s}^{-1}(K)$. Then $y = \phi_{t,s}(x) \in K$. Setting $y_n := \phi_{t,s}^n(x)$, we have $\lim_{n \to \infty} y_n = y$ by Lemma 2.2, which means that, for all $r > 0$, there exists $n \in \mathbb{N}$ such that, for all $k \geq n$, $\phi_{t,s}^k(x) \in B_r(y) \subset K_r$. This proves the $\subset$ direction of (2.13).

Now suppose x belongs to the right-hand side of (2.13). Then, for all $r > 0$, there exists $n \in \mathbb{N}$ such that $x \in (\phi_{t,s}^k)^{-1}(K_r)$ for all $k \geq n$. Set $y_k := \phi_{t,s}^k(x)$, so that we have $y_k \subset K_r$ for all $k \geq n$. We have $y := \lim_{k \to \infty} y_k = \lim_{k \to \infty} \phi_{t,s}^k(x) = \phi_{t,s}(x)$ by Lemma 2.2. On the other hand, we also have $y \in \overline{K_r}$, and so

$$\phi_{t,s}(x) \subset \bigcap_{r>0} \overline{K_r} = K.$$

Thus, the $\supset$ direction of (2.13) is established.

---

[5] That is, we abuse notation and suppose that $b^n = b^{\varepsilon_n}$ for $(b^\varepsilon)_{\varepsilon>0}$ satisfying (2.3) and some $(\varepsilon_n)_{n \in \mathbb{N}}$ satisfying $\lim_{n \to \infty} \varepsilon_n = 0$.

The continuity of $\phi_{t,s}$ and the compactness of K imply that $\phi_{t,s}^{-1}(K)$ is closed. We note also that $(\phi_{t,s}^k)^{-1} = \phi_{s,t}^k$ satisfies (2.5) uniformly in k, because the bound only depends on the constant C_0 in the linear growth bound of (2.1), which is also satisfied by $-b^k$. This along with (2.13) implies that $\phi_{t,s}^{-1}(K)$ is bounded, and thus compact.

We now show that $\phi_{t,s}$ is surjective. Using again the bound (2.5) satisfied uniformly in k for $\phi_{s,t}^k$, we set $x_n := (\phi_{t,s}^n)^{-1}(y) = \phi_{s,t}^n(y)$ and note that $(x_n)_{n \in \mathbb{N}}$ is bounded. Passing to a subsequence, we have $\lim_{k \to \infty} x_{n_k} = x$ for some $x \in \mathbb{R}^d$, and then $y = \phi_{t,s}^{n_k}(x_k)$, so that $y = \lim_{k \to \infty} \phi_{t,s}^{n_k}(x_k) = \phi_{t,s}(x)$.

Finally, we show $\phi_{t,s}^{-1}(K)$ is connected. For each $k \in \mathbb{N}$, $(\phi_{t,s}^k)^{-1}(K_r)$ is connected, and therefore so is the intersection $\bigcap_{k \geq n}(\phi_{t,s}^k)^{-1}(K_r)$ for each n. These sets are nested in n, so taking the union in $n \in \mathbb{N}$ yields a connected set. Taking the intersection over $r > 0$ gives the connectedness of $\phi_{t,s}^{-1}(K)$. $\square$

Remark 2.5. The fact that the approximate backward flows converge *uniformly* to $\phi_{t,s}$ is used in the second-to-last paragraph of the proof, in order to show that $\phi_{t,s}$ is surjective.

Proof of Proposition 2.2. We identify the set by

$$A_{t,s} = \Big\{ y \in \mathbb{R}^d : \text{ there exists } x \in \phi_{t,s}^{-1}(\{y\}) \text{ such that}$$

$$\phi_{t,s} \text{ is differentiable at } x \text{ and } J_{t,s}(x) \neq 0 \Big\}.$$

We first check that $A_{t,s}$ has full measure. Its complement consists of

$$\mathbb{R}^d \backslash A_{ts} = \Big\{ y \in \mathbb{R}^d : J_{t,s} = 0 \text{ at the points of differentiability of } \phi_{t,s} \text{ on } \phi_{t,s}^{-1}(\{y\}) \Big\}$$

$$\cup \Big\{ y \in \mathbb{R}^d : \phi_{t,s} \text{ is not differentiable anywhere in } \phi_{t,s}^{-1}(\{y\}) \Big\}.$$

The fact that $\phi_{t,s}$ is differentiable a.e. and the change of variables formula then give

$$|\mathbb{R}^d \backslash A_{t,s}| = \int_{\mathbb{R}^d} \mathbf{1}\{\phi_{t,s}(x) \in \mathbb{R}^d \backslash A_{t,s}\} J_{t,s}(x) dx = 0.$$

It remains to show that $\phi_{t,s}^{-1}(\{y\})$ is a singleton for all $y \in A_{t,s}$. By Lemma 2.4, $\phi_{t,s}^{-1}(\{y\})$ is nonempty, compact, and connected. Suppose $x, \tilde{x} \in \phi_{t,s}^{-1}(\{y\})$ are such that $J_{t,s}(x) \neq 0$. A Taylor expansion gives

$$y = \phi_{t,s}(\tilde{x}) = \phi_{t,s}(x) + \nabla_x \phi_{t,s}(x) \cdot (x - \tilde{x}) + o(|x - \tilde{x}|)$$
$$= y + \nabla_x \phi_{t,s}(x) \cdot (x - \tilde{x}) + o(|x - \tilde{x}|).$$

The invertibility of $\nabla_x \phi_{t,s}(x)$ then implies that, if $|\tilde{x} - x|$ is sufficiently small, then $\tilde{x} = x$, or, in other words, x is an isolated point. But then the connected set $\phi_{t,s}(\{y\})$ must be equal to $\{x\}$, and we call $x = \phi_{s,t}(y)$.

For $y \in A_{t,s}$, we then have $(\phi_{t,s} \circ \phi_{s,t})(y) = y$. Since $\phi_{t,s}^{-1}(\{y\})$ is nonempty for any $y \in \mathbb{R}^d$, we may define a version of $\phi_{s,t}$ on all of $\mathbb{R}^d$ by imposing that

$\phi_{s,t}(y) \in (\phi_{t,s}^{-1})(\{y\})$ for any $y \in A_{t,s}$. For this version, we have $\phi_{t,s} \circ \phi_{s,t} = \mathrm{Id}$ everywhere on $\mathbb{R}^d$. Suppose now that $y \in A_{t,s}$ and $\lim_{n \to \infty} y_n = y$ for some sequence $(y_n)_{n \in \mathbb{N}} \subset \mathbb{R}^d$. Then

$$\lim_{n \to \infty} (\phi_{t,s} \circ \phi_{s,t})(y_n) = (\phi_{t,s} \circ \phi_{s,t})(y).$$

We have

$$(\phi_{s,t}(y_n))_{n \in \mathbb{N}} \subset \bigcup_{n \in \mathbb{N}} (\phi_{t,s})^{-1}(\{y_n\}),$$

which implies by Lemma 2.4 that $(\phi_{s,t}(y_n))_{n \in \mathbb{N}}$ is bounded. Letting z be a limit point of this set, we have, by continuity of the backward flow, that $y = \lim_{n \to \infty} y_n = \phi_{t,s}(z)$, and therefore $z = \phi_{s,t}(y)$. $\qquad \square$

Remark 2.6. We shall see in Sect. 4 that the forward flow is always BV in space. Therefore, the "forward Jacobian" $J_{t,s}$ for $t > s$ can only be understood as a measure. Indeed, returning to the example $b(t, x) = \mathrm{sgn}\, x$ on $\mathbb{R}$, the right inverse $\phi_{t,s}$ of $\phi_{s,t}$ given by (2.7) is $\phi_{t,s}(x) = x + (\mathrm{sgn} x)(t - s)$ for $s \leq t$, which is discontinuous only at 0. The backward Jacobian is given by $J_{s,t}(x) = \mathbf{1}\{|x| \geq t - s\}$, and the forward one is $J_{t,s} = 1 + 2(t - s)\delta_0$.

Remark 2.7. The formula $\phi_{s,t} \circ \phi_{t,s} = \mathrm{Id}$ makes sense a.e. if $s < t$, because $\phi_{s,t}$ is Lipschitz and $\phi_{t,s}$ is measurable. On the other hand, $\phi_{t,s}$ is not also a left-inverse, since the formula $\phi_{t,s} \circ \phi_{s,t}$ does not make sense. In the above example, $\phi_{s,t}(x)$ is equal to 0, for $|x| \leq t - s$, and 0 is a point of discontinuity for $\phi_{t,s}$. In general, the concentration of $\phi_{s,t}$ on sets of measure 0 forbids applying $\phi_{t,s}$ as a left-inverse.

2.4. Compressive Stochastic Flows

We now fix a matrix-valued map

$$\Sigma \in L^2([0, T], C^{0,1}(\mathbb{R}^d; \mathbb{R}^{d \times m})), \tag{2.14}$$

and assume that

$$W : \Omega \times [0, T] \to \mathbb{R}^m \quad \text{is a standard Brownian motion}$$
$$\text{on a given probability space } (\Omega, \mathcal{F}, \mathbb{P}, \mathbb{E}). \tag{2.15}$$

In order to extend the results in the preceding subsections, and, in particular, to bypass the difficulties of the backward time direction, we consider forward SDEs with drift satisfying the opposite of (2.1), that is,

$$B : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d, \quad -B \text{ satisfies (2.1)}, \tag{2.16}$$

and consider the flow

$$\begin{cases} d_s \Phi_{s,t}(x) = B(s, \Phi_{s,t}(x))ds + \Sigma(s, \Phi_{s,t}(x))dW_s, \ s \in [t, T], \\ \Phi_{t,t}(x) = x. \end{cases} \tag{2.17}$$

Once again, (2.17) must be understood in the Filippov sense, which means, for $s \in [t, T]$,

$$\Phi_{s,t}(x) = x + \int_t^s \alpha_r \, dr + \int_t^s \Sigma(r, \Phi_{r,t}(x)) dW_r, \quad \alpha_s \in B(s, \Phi_{s,t}(x)), \quad (2.18)$$

and we remark that our assumptions will allow us to always consider probabilistically strong solutions; that is, we solve (2.18) path by path for almost every continuous W with respect to the Wiener measure. Depending on the context in later sections (in particular, the time direction of solvability for the transport and continuity equations), we consider different examples for B and Σ for which these assumptions are satisfied.

Lemma 2.5. *For every* $(t, x) \in [0, T] \times \mathbb{R}^d$ *and* $\mathbb{P}$-*almost surely, there exists a unique strong solution* $\Phi_{s,t}(x)$ *of* (2.17) *defined on* $[t, T] \times \mathbb{R}^d$. *Moreover, for all* $p \in [2, \infty)$, *there exists a constant* $C = C_p > 0$ *depending only on the assumptions* (2.1) *and* (2.14) *such that*

$$\mathbb{E}|\Phi_{s,t}(x) - \Phi_{s,t}(y)|^p \le C|x - y|^p \quad \text{for all } 0 \le t \le s \le T \text{ and } x, y \in \mathbb{R}^d,$$
(2.19)

$$\mathbb{E}|\Phi_{s,t}(x)|^p \le C(|x|^p + 1) \quad \text{for all } -0 \le t \le s \le T \text{ and } x \in \mathbb{R}^d,$$
(2.20)

and

$$\begin{cases} \mathbb{E}|\Phi_{s_1,t}(x) - \Phi_{s_2,t}(x)|^p \le C(1 + |x|)|s_1 - s_2|^{p/2} \\ \text{for all } t \in [0, T], s_1, s_2 \in [t, T], \text{ and } x \in \mathbb{R}^d. \end{cases}$$
(2.21)

With probability one, for all $0 \le r \le s \le t \le T$, $\Phi_{t,s} \circ \Phi_{s,r} = \Phi_{t,r}$. *If* $(b^\varepsilon)_{\varepsilon>0}$ *are regularizations satisfying* (2.3), *then, with probability one, the corresponding stochastic flows* Φ^ε *converge locally uniformly as* $\varepsilon \to 0$ *to* Φ.

Proof. For $\varepsilon > 0$, let B^ε be the convolution of B in space by a standard mollifier (so that $b^\varepsilon := -B^\varepsilon$ satisfies (2.3)), and let $\Phi_{t,s}^\varepsilon$ denote the corresponding stochastic flow. Itô's formula, the one-sided Lipschitz assumption, and the Lipschitz continuity of Σ yield, for any $p \ge 2$ and some $C \in L_+^1([0, T])$,

$$\frac{\partial}{\partial t}\mathbb{E}|\Phi_{t,s}^\varepsilon(x) - \Phi_{t,s}^\varepsilon(y)|^p \le C(t)\mathbb{E}|\Phi_{t,s}^\varepsilon(x) - \Phi_{t,s}^\varepsilon(y)|^p,$$

which, along with Grönwall's inequality, leads to the first statement. The other two estimates are proved similarly, with constants independent of $\varepsilon > 0$.

In view of (2.19) and (2.21), the Kolmogorov continuity criterion then yields, for any $R > 0$, $p \ge 2$ and $\delta \in (0, 1)$, a constant $C = C_{R,p,\delta} > 0$ such that, for all $s \in [0, T]$, $\lambda \ge 1$ and $\varepsilon > 0$,

$$\mathbb{P}\left(\sup_{x,y \in B_R} \sup_{r,s \in [s,T]} \frac{|\Phi_{t,s}^\varepsilon(x) - \Phi_{r,s}^\varepsilon(y)|}{|x - y|^{1-\delta} + |t - s|^{\frac{1}{2}(1-\delta)}} > \lambda\right) \le \frac{C}{\lambda^p}.$$

It follow that the probability measures on $C([s, T] \times \mathbb{R}^d; \mathbb{R}^d)$ induced by the random variables $(\Phi^\varepsilon_{\cdot,s})_{\varepsilon>0}$ are tight with respect to the topology of locally uniform convergence, and therefore converge weakly along a subsequence as $\varepsilon \to 0$ to a probability measure that gives rise to a weak (in the probabilistic sense) solution of (2.17), for which the estimates in the statement of the lemma continue to hold.

A similar computation to the one above reveals that, for a fixed probability space and almost every Brownian path W, the solution of (2.17) is unique. The pathwise uniqueness then implies, by a standard argument due to Yamada and Watanabe [66], that there is a unique strong solution for every $x \in \mathbb{R}^d$. □

Remark 2.8. It is an open question whether $\Phi_{t,s}$ is Lipschitz continuous, even if B is Lipschitz. When B is Lipschitz and $\Sigma \in C^{1,\alpha}$ for some $\alpha \in (0, 1]$, it turns out the flow $\Phi_{t,s}$ is $C^{1,\alpha'}$ for any $\alpha' \in (0, \alpha)$, but it is not clear how to extend this to the case where $-B$ satisfies the one-sided Lipschitz bound from below.

As a consequence, an understanding of the Jacobian $\det(\nabla_x \Phi_{t,s}(x))$, or of the stability with respect to regularizations of B, is considerably more complicated in the stochastic case. The results of Sect. 4, where we discuss the expansive regime, are therefore constrained to the first-order case, and we relegate the second-order analysis to future work. One exception is when Σ is independent of the spatial variable, in which case a change of variables relates the SDE to an ODE of the form (2.2) with a random b.

2.5. Small Noise Approximations

We return to the backward flow $\phi_{t,s}$, $0 \le t \le s \le T$, from Lemma 2.2. Recall that the backward flow also corresponds to the forward flow for $-b$; that is,

$$\frac{\partial}{\partial s}\phi_{t,s}(x) = -b(s, \phi_{t,s}(x)), \quad s \ge t, \quad \phi_{t,t}(x) = x. \tag{2.22}$$

For $\varepsilon > 0$, let $\phi^\varepsilon_{t,s}(x)$ denote the following stochastic flow

$$d_s\phi^\varepsilon_{t,s}(x) = -b(s, \phi^\varepsilon_{t,s}(x))ds + \varepsilon dW_s \quad s \ge t, \quad \phi^\varepsilon_{t,t}(x) = x, \tag{2.23}$$

where W is now a d-dimensional Brownian motion. We note that (2.17) falls under the assumptions of Lemma 2.5, but in fact (2.23) admits a unique strong solution as soon as b is merely locally bounded [39,65]. In general, the limiting solutions as $\varepsilon \to 0$ are not unique; however, we immediately have the following as a consequence of Lemma 2.5.

Proposition 2.3. *For every $\varepsilon > 0$, there exists a unique strong solution of (2.23). Moreover, as $\varepsilon \to 0$, ϕ^ε converges locally uniformly to ϕ.*

If $J^\varepsilon = \det(\nabla_x\phi^\varepsilon)$, then, as $\varepsilon \to 0$, J^ε converges weak-$\star$ in $L^\infty([s, T] \times \mathbb{R}^d)$ and weakly in $C([s, T], L^1_{loc}(\mathbb{R}^d))$ to J.

2.6. Some Bibliographical Remarks

We conclude this section by placing the above results in the context of the existing literature. As has been mentioned, the well-posedness and properties of the backwards flow in Sect. 2.1 are well studied, and date back to at least the work of Filippov [42]. For the particular properties of the Jacobian stated in subsection 2.2, especially the weak-$\star$ convergence in L^∞, we expand on arguments from the Appendix of [25]. Our result on strong convergence when $d = 1$ relies on arguments as for Aubin-Lions compactness Lemmas [18,54].

Meanwhile, the forward flow in Sect. 4.3 is comparatively less studied. Our approach to uniquely identifying the forward flow as the right-inverse of the backward flow is similar to that in the appendix of [29], with the argument expanded so as to prove the almost-everywhere continuity, which was not previously known. As mentioned, we further expand the properties of the forward flow in Sect. 4 below.

As in the ODE case, the theory for compressive SDE flows such as those considered in Sect. 2.4 is well-understood. For instance, in [62], this situation is studied in a still more general setting (the constant C_1 in (2.1) is allowed to depend additionally on x and y, with some integrability assumptions), and, similarly as in Sect. 3.2 below, this allows for a theory of measure-valued solutions of the Fokker–Planck equation (3.4). By contrast, degenerate stochastic regular Lagrangian flows and degenerate Fokker-Planck equations with $(\operatorname{div} b)_- \in L^\infty$ and $\operatorname{div} b \notin L^1$ are far less studied, and, as far as we know, our results in Sect. 4.5 below (which rely on the properties of compressive flows established in this section) are the first in this direction.

3. The Compressive Regime

In this section, we consider the transport and continuity equations in the so-called compressive regime. That is, for velocity field b satisfying (2.1), we study the TVP for the nonconservative equation

$$-\frac{\partial u}{\partial t} + b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d \quad \text{and} \quad u(T, \cdot) = u_T \quad \text{in } \mathbb{R}^d, \quad (3.1)$$

and the IVP for the conservative equation

$$\frac{\partial f}{\partial t} - \operatorname{div}(b(t, x) f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d \quad \text{and} \quad f(0, \cdot) = f_0. \quad (3.2)$$

We recall that $\operatorname{div} b$ is bounded from below, and therefore, the direction of time for (3.1) and (3.2) does not allow for a solution theory in Lebesgue spaces, due to the concentrative nature of the backward flow analyzed in the previous section. The TVP (3.1) will be solved in the space of continuous functions, while (3.2) is solved in the dual space of Radon measures.

We also obtain analogous results for the second-order equations

$$-\frac{\partial u}{\partial t} - \operatorname{tr}[a(t, x)\nabla^2 u] + b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d \quad \text{and} \quad u(T, \cdot) = u_T$$

$$(3.3)$$

and

$$\frac{\partial f}{\partial t} - \operatorname{div}\left[\operatorname{div}(a(t,x)f) - b(t,x)f\right] = 0 \quad \text{in } (0,T) \times \mathbb{R}^d \quad \text{and} \quad f(0,\cdot) = f_0, \tag{3.4}$$

where $a = \frac{1}{2}\sigma\sigma^T$ for $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$, satisfying that

$$\sup_{x \in \mathbb{R}^d} \frac{|\sigma(\cdot,x)|}{1+|x|} + \sup_{y,z \in \mathbb{R}^d} \frac{|\sigma(\cdot,y) - \sigma(\cdot,z)|}{|y-z|} \in L^2([0,T]). \tag{3.5}$$

3.1. The Nonconservative Equation

3.1.1. Representation Formula   When interpreting (3.1) in the distributional sense, we are constrained to seek solutions that are continuous. Indeed, the distribution

$$b \cdot \nabla u = \operatorname{div}(bu) - (\operatorname{div} b)u$$

pairs the solution u with div b, which is a measure in general. The other motivating factor is the formal representation formula for the solution of the TVP (3.1), which is given in terms of the backward flow:

$$u(t,x) = u_T(\phi_{t,T}(x)) \quad \text{for } (t,x) \in [0,T] \times \mathbb{R}^d. \tag{3.6}$$

This formula and the Lipschitz continuity of $\phi_{t,T}$ given in Lemma 2.2 suggest that the solution operator for (3.1) should preserve continuity. In fact, the formula (3.6) defines a distributional solution, which is uniquely obtained from limits of natural regularizations of the equation.

Theorem 3.1. *If $u_T \in C(\mathbb{R}^d)$, then the function u in (3.6) is a distributional solution of (3.1). Moreover, if $(b^\varepsilon)_{\varepsilon>0}$ satisfy (2.3) and u^ε is the corresponding solution of (3.1) with velocity field b^ε, then, as $\varepsilon \to 0$, u^ε converges locally uniformly to u.*

Proof. The unique solution u^ε for the regularized velocity field is given by $u^\varepsilon(t,\cdot) = u_T \circ \phi^\varepsilon_{t,T}$, where ϕ^ε is the flow corresponding to b^ε. By Lemma 2.2, as $\varepsilon \to 0$, ϕ^ε converges locally uniformly to ϕ, and so the local-uniform convergence to u follows from the continuity of u_0.

Multiplying the equation for u^ε by some $\psi \in C^1_c((0,T) \times \mathbb{R}^d)$ and integrating by parts gives

$$\int_0^T \int_{\mathbb{R}^d} u^\varepsilon(t,x) \left(\partial_t \psi(t,x) - b^\varepsilon(t,x) \cdot \nabla \psi(t,x) + (\operatorname{div} b^\varepsilon(t,x))\psi(t,x) \right) dx\, dt = 0.$$

As $\varepsilon \to 0$, $b^\varepsilon \to b$ almost everywhere and div $b^\varepsilon \rightharpoonup$ div b weakly in the sense of measures, and so the fact that u is a distributional solution follows. $\square$

Turning next to the second-order equation (3.3), we identify a solution candidate with the appropriate stochastic flow. We do so by changing the time direction in b and σ and considering the SDE

$$\begin{cases} d_s \Phi_{s,t}(x) = -b(s, \Phi_{s,t}(x))ds + \sigma(s, \Phi_{s,t}(x))dW_s, \ s \in [t, T], \\ \Phi_{t,t} = \text{Id}, \end{cases} \quad (3.7)$$

where W is as in (2.15). Note that (3.7) is of the type in (2.17) and thus falls within the assumptions of Lemma 2.5. In particular, if u_T is continuous, then, in view of (2.19)-(2.21), the formula

$$u(t, x) = \mathbb{E}[u_T(\Phi_{T,t}(x))] \quad (3.8)$$

defines a continuous function. Moreover, if u_T is Lipschitz, then $u(t, \cdot)$ is Lipschitz for all $t > 0$, and $1/2$-Hölder continuous in time. Note that, in this case, the distribution $\text{tr}[a\nabla^2 u] = \text{div}(a\nabla u) - (\text{div } a) \cdot \nabla u$ makes sense, because ∇u and $\text{div } a$ both belong to L^∞.

The following is proved exactly as for Theorem 3.1, with the use of the estimates in Lemma 2.5.

Theorem 3.2. *Let $u_T \in C(\mathbb{R}^d)$ be uniformly continuous and define u by (3.8). If $(b^\varepsilon)_{\varepsilon>0}$ satisfy (2.3) and u^ε is the corresponding solution of (3.3) with velocity b^ε, then, as $\varepsilon \to 0$, u^ε converges locally uniformly to u. Moreover, if $u_T \in C^{0,1}$, then*

$$\sup_{(t,x,y)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d} \frac{|u(t, x) - u(t, y)|}{|x - y|} + \sup_{(r,s,z)\in[0,T]\times\mathbb{R}^d} \frac{|u(r, z) - u(s, z)|}{|r - s|^{1/2}(1 + |z|)} < \infty,$$

and u is a distributional solution of (3.3).

As a special case, we consider, for $\varepsilon > 0$, the "viscous" version of (3.1), that is

$$-\partial_t u^\varepsilon - \frac{\varepsilon^2}{2}\Delta u^\varepsilon + b(t, x) \cdot \nabla u^\varepsilon = 0 \ \text{ in } (0, T) \times \mathbb{R}^d, \quad u^\varepsilon(T, \cdot) = u_T. \quad (3.9)$$

This uniformly parabolic equation has a unique classical solution for any uniformly continuous $u_T : \mathbb{R}^d \to \mathbb{R}$, which, moreover, is given by $u^\varepsilon(t, x) = \mathbb{E}[u_T(\phi_{t,T}^\varepsilon(x))]$, where now ϕ^ε denotes the solution of the SDE (2.23) from the previous section. Arguing just as in Theorem 3.1 and invoking Proposition 2.3 immediately gives the following:

Theorem 3.3. *As $\varepsilon \to 0$, the solution u^ε converges locally uniformly to the function u given by (3.6).*

3.1.2. Viscosity Solutions   Although (3.6) and (3.8) are the distributional solutions that arise uniquely through regularization (either of b or through vanishing viscosity limits), it turns out that distributional solutions are not unique in general (see Sect. 3.1.3 below). It is then a natural question as to whether the "good" solutions can be characterized other than as limits of regularizations, or by the explicit formulae. For example, this is done for the one-dimensional problem in [59] by introducing a sort of entropy condition.

We give a different characterization here using the theory of viscosity solutions [35], which covers both the first- and second-order problems. We present the results here only in the second-order case, which includes the first-order equations when $a = 0$.

We define, for $(t, x, p) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$,

$$\underline{b}(t, x, p) = \liminf_{z \to x} b(t, z) \cdot p \quad \text{and} \quad \overline{b}(t, x, p) = \limsup_{z \to y} b(t, z) \cdot p.$$

For fixed $(t, x) \in [0, T] \times \mathbb{R}^d$, $\underline{b}(t, x, \cdot)$ and $\overline{b}(t, x, \cdot)$ are Lipschitz continuous on $\mathbb{R}^d$, and, for fixed $(t, p) \in [0, T]$, $\underline{b}(t, \cdot, p)$ and $\overline{b}(t, \cdot, p)$ are respectively lower and upper semicontinuous.

The following definition of viscosity (sup, super) solutions closely resembles the one in [55]:

Definition 3.1. An upper-semicontinuous (resp. lower-semicontinuous) function u is called a subsolution (resp. supersolution) of (3.3) if, for all $\psi : [0, T] \times \mathbb{R}^d$ that are C^1 in t and C^2 in x, it holds that

$$-\frac{d}{dt} \max_{x \in \mathbb{R}^d} \{u(t, x) - \psi(t, x)\}$$

$$\leq \inf \left\{ \text{tr}[a(t, y)\nabla^2 \psi(t, y)] - \underline{b}(t, y, \nabla \psi(t, y)) : y \in \arg\max\{u(t, \cdot) - \psi(t, \cdot)\} \right\}$$

(resp.

$$-\frac{d}{dt} \min_{x \in \mathbb{R}^d} \{u(t, x) - \psi(t, x)\}$$

$$\geq \sup \left\{ \text{tr}[a(t, y)\nabla^2 \psi(t, y)] - \overline{b}(t, y, \nabla \psi(t, y)) : y \in \arg\min\{u(t, \cdot) - \psi(t, \cdot)\} \right\} \Big).$$

If $u \in C([0, T] \times \mathbb{R}^d)$ is both a sub and supersolution, we say u is a solution.

The comparison principle is proved by doubling the space variable. In particular, we have the following lemma, which follows exactly by methods as in [34,57,58]. For $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$, we define the nonnegative matrix

$$A(t, x, y) := \begin{pmatrix} \sigma(t, x) \\ \sigma(t, y) \end{pmatrix} \begin{pmatrix} \sigma(t, x)^T & \sigma(t, y)^T \end{pmatrix}.$$

Lemma 3.1. *Assume u and v are respectively a sub and supersolution of* (3.3). *Then $w(t, x, y) = u(t, x) - v(t, y)$ is a subsolution of*

$$-\partial_t w - \mathrm{tr}[a(t, x, y)\nabla^2_{(x,y)} w] + \underline{b}(t, x, \nabla_x w) - \overline{b}(t, x, -\nabla_y w) \le 0.$$

We may now state and prove the comparison principle.

Theorem 3.4. *If u and v are respectively a sub and supersolution of* (3.3) *such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \frac{u(t, x)}{1 + |x|} + \sup_{(s,y)\in[0,T]\times\mathbb{R}^d} \frac{-v(s, y)}{1 + |y|} < \infty,$$

then $t \mapsto \sup_{x\in\mathbb{R}^d} \{u(t, x) - v(t, x)\}$ is nondecreasing.

Proof. Define $w(t, x, y) := u(t, x) - v(t, y)$, fix $\delta, \varepsilon > 0$, and define $\Phi_{\delta,\varepsilon}(x, y) = \frac{1}{2\delta}|x - y|^2 + \frac{1}{2\varepsilon}(|x|^2 + |y|^2)$. In view of the growth of u and v in x, for all $t \in [0, T]$, the map $w(t, \cdot, \cdot) - \Phi_{\delta,\varepsilon}(x, y)$ attains a maximum on $\mathbb{R}^d \times \mathbb{R}^d$. Moreover, standard arguments from the theory of viscosity solutions (see for instance [35, Lemma 3.1]) imply that there exist $\rho_\delta > 0$ and λ_ε such that $\lim_{\delta\to 0} \rho_\delta^2/\delta = \lim_{\varepsilon\to 0} \varepsilon\lambda_\varepsilon^2 = 0$, and

$$|x - y| \le \rho_\delta \quad \text{and} \quad |x| + |y| \le \lambda_\varepsilon \quad \text{for all } (x, y) \in \arg\max\{w(t, \cdot, \cdot) - \Phi_{\delta,\varepsilon}\}, \quad t \in [0, T].$$

Therefore, if $t \in [0, T]$ and $(x, y) \in \arg\max\{w(t, \cdot, \cdot) - \Phi_{\delta,\varepsilon}\}$, we have, for some $C \in L^1_+([0, T])$,

$$\mathrm{tr}[a(t, x, y)\nabla^2_{(x,y)}\Phi_{\delta,\varepsilon}(x, y)]$$
$$= \mathrm{tr}\left[\left(\frac{1}{\delta}\begin{pmatrix} \mathrm{Id} & -\mathrm{Id} \\ -\mathrm{Id} & \mathrm{Id} \end{pmatrix} + \varepsilon\begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} \end{pmatrix}\right)\begin{pmatrix} \sigma(t, x) \\ \sigma(t, y) \end{pmatrix}\left(\sigma(t, x)^T \ \sigma(t, y)^T\right)\right]$$
$$\le C(t)\left(\frac{\rho_\delta^2}{\delta} + \varepsilon\lambda_\varepsilon^2\right)$$

and

$$-\underline{b}\left(t, x, \nabla_x \Phi_{\delta,\varepsilon}(x, y)\right) + \overline{b}\left(t, y, -\nabla_y \Phi_{\delta,\varepsilon}(x, y)\right)$$
$$= \limsup_{(z,w)\to(x,y)}\left\{-b(t, z)\cdot\left(\frac{x - y}{\delta} + \beta x\right) + b(t, w)\cdot\left(\frac{x - y}{\delta} - \beta y\right)\right\}$$
$$= \limsup_{(z,w)\to(x,y)}\left\{-(b(t, z) - b(t, w))\cdot\frac{z - w}{\delta} - b(t, z)\cdot\beta z + b(t, w)\cdot\beta w\right\}$$
$$\le C(t)\left(\frac{\rho_\delta^2}{\delta} + \varepsilon + \varepsilon\lambda_\varepsilon\right).$$

It now follows from Definition 3.1 and Lemma 3.1 that, for some $C_{\delta,\varepsilon} \in L^1_+([0, T])$ satisfying $\lim_{(\delta,\varepsilon)\to(0,0)} C_{\delta,\varepsilon} = 0$ in $L^1([0, T])$,

$$t \mapsto \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d}\{w(t, x, y) - \Phi_{\delta,\varepsilon}(x, y)\} - \int_t^T C_{\delta,\varepsilon}(s)ds$$

is nondecreasing. The result follows upon sending δ and ε to 0. $\qquad\square$

As a consequence of the comparison theorem, the "good" distributional solution of (3.3) can be uniquely characterized.

Theorem 3.5. *Assume* $u_T : \mathbb{R}^d \to \mathbb{R}$ *is uniformly continuous and* $u_T \cdot (1+|\cdot|^{-1}) \in L^\infty$. *Then* (3.8) *is the unique viscosity solution of* (3.3).

Proof. The fact that (3.8) defines a viscosity solution is due to Theorem 3.2 and the stability properties of viscosity solutions.[6] In view of Lemma 2.5 and the growth of u_T, we may appeal to Theorem 3.4 to conclude that (3.8) is the only viscosity solution of the terminal value problem (3.3).   □

3.1.3. (Non)equivalence of Distributional and Viscosity Solutions   For $x \in \mathbb{R}$, set $b(t, x) = \text{sgn } x$ and $u_T(x) = |x|$. Using the formula (2.7) for the backward flow, the solution (3.6) becomes

$$u(t, x) = (|x| - (T - t))_+. \tag{3.10}$$

However, the Lipschitz function

$$v(t, x) = |x| - (T - t) \tag{3.11}$$

is another distributional solution (and in fact satisfies the equation a.e.). It can also be checked directly that (3.11) does not give a viscosity solution of (3.1). Indeed, note that $v(t, x) - t$ attains a global minimum at any $(t, 0) \in [0, T] \times \mathbb{R}$. Applying the supersolution definition with $\phi(t, x) = t$ yields the contradictory $-1 \geq 0$.

The uniqueness of distributional solutions fails even if b is continuous. Indeed, if $0 < \alpha < 1$ and $b(t, x) = \text{sgn } x |x|^\alpha$ and $u_T(x) = |x|^{1-\alpha}$, then, arguing similarly as in the above example,

$$u(t, x) = \left(|x|^{1-\alpha} - (1 - \alpha)(T - t) \right)_+ \tag{3.12}$$

and

$$v(t, x) = |x|^{1-\alpha} - (1 - \alpha)(T - t) \tag{3.13}$$

are two distributional solutions, and (3.12) is the one corresponding to (3.6). Once again, (3.13) can directly be seen to fail the viscosity supersolution property.

In the first example above, u_T is Lipschitz while b is discontinuous, and, while b is continuous in the second example, we take u_T to be non-Lipschitz. This should be compared with the following sufficient criterion for equivalence.

Theorem 3.6. *If* $b \in C([0, T] \times \mathbb{R}^d)$ *satisfies* (2.1) *and* $u_T \in C^{0,1}(\mathbb{R}^d)$, *then there exists a unique distributional solution* $u \in C([0, T], C^{0,1}(\mathbb{R}^d))$, *and it is given by* (3.6).

---

[6] Note that smooth solutions of the equation corresponding to b^ε satisfying (2.3), or of the viscous equation (3.9), are viscosity solutions in the sense of Definition 3.1.

Proof. Let $\rho \in C_c^\infty$ be a standard mollifier and, for $\varepsilon > 0$, set $\rho_\varepsilon(x) = \varepsilon^{-d}\rho(\varepsilon^{-1}x)$. Let $u \in C([0, T], C^{0,1}(\mathbb{R}^d))$ be a distributional solution of (3.1) and define $u_\varepsilon = u * \rho_\varepsilon$. Then

$$-\partial_t u_\varepsilon + b \cdot \nabla u_\varepsilon = r_\varepsilon \quad \text{in } (0, T) \times \mathbb{R}^d, \tag{3.14}$$

where

$$r_\varepsilon(t, x) = \int_{\mathbb{R}^d} (b(t, y) - b(t, x)) \cdot \nabla u(t, y)\rho_\varepsilon(x - y)dy.$$

Note that $r_\varepsilon \in C([0, T] \times \mathbb{R}^d)$, and u_ε solves (3.14) in the sense of viscosity solutions. Moreover, the continuity of b and boundedness of ∇u imply that $r_\varepsilon \xrightarrow{\varepsilon \to 0} 0$ locally uniformly. Standard stability results from the theory of viscosity solutions then imply that the limit u of u_ε is the unique viscosity solution of (3.1). $\quad\square$

The above result can be extended by studying the interplay between regularity of b and u.

Theorem 3.7. *Suppose that $\alpha, \beta \in (0, 1]$ satisfy $\alpha + \beta > 1$, b satisfies (2.1) and $\sup_{t\in[0,T]}[b(t, \cdot)]_{C^\alpha} < \infty$, and u is a distributional solution of (3.1) such that $\sup_{t\in[0,T]}[u(t, \cdot)]_{C^\beta} < \infty$. Then u is the unique viscosity solution of (3.1).*

Remark 3.1. The condition on $\alpha + \beta$, and, in particular, the strict inequality, is sharp, as the example above with $b(x) = \text{sgn } x|x|^\alpha$ and $u_T(x) = |x|^{1-\alpha}$ shows.

Proof of Theorem 3.7. Arguing similarly as for Theorem 3.6, it suffices to prove that

$$r_\varepsilon = (b \cdot \nabla u) * \rho_\varepsilon - b \cdot \nabla(u * \rho_\varepsilon) \xrightarrow{\varepsilon \to 0} 0 \quad \text{locally uniformly,}$$

where ρ_ε is a standard mollifier. We note that $r_\varepsilon = M_\varepsilon[b(t, \cdot), u(t, \cdot)]$, where the bilinear operator M_ε is defined, for sufficiently regular $(B, U) : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}$, by

$$M_\varepsilon[B, U] = \int_{\mathbb{R}^d} (B(y) - B(x)) \cdot \nabla U(y)\rho_\varepsilon(x - y)dy.$$

Standard interpolation arguments give, for some $C > 0$ depending on α and β, for all $(B, U) \in C^\alpha \times C^\beta$,

$$|M_\varepsilon[B, U]| \le C\varepsilon^{\alpha+\beta-1}[B]_{C^\alpha}[U]_{C^\beta}.$$

Therefore $|r_\varepsilon(t, x)| \le C[b(t, \cdot)]_{C^\alpha}[u(t, \cdot)]_{C^\beta}\varepsilon^{\alpha+\beta-1}$, and we conclude upon sending $\varepsilon \to 0$. $\quad\square$

3.2. The Conservative Equation

3.2.1. Duality Solutions   For either of the two conservative equations (3.2) and (3.4), the tendency of the backward flow to concentrate on sets of Lebesgue measure zero implies that, even if f_0 is absolutely continuous with respect to the Lebesgue measure, $f(t, \cdot)$ may develop a singular part for $t > 0$.

This presents an obstacle in defining solutions in the sense of distributions, since the product of the discontinuous vector field b with a singular measure f may not be well-defined. This same issue arose in the works [22,25], and also in studying the the nonconservative equation (4.2) in Lebesgue spaces (see Sect. 4 below). The approach in these works was to define solutions through duality with the dual equation, for which particular distributional solutions could be defined in a stable, unique way. In this compressive regime, we do the same for (3.2), and directly define solutions in duality with the nonconservative equation.

Definition 3.2. A map $f \in C([0, T], \mathcal{M}_{\text{loc,w}})$ is called a solution of (3.2) if, for all $t \in [0, T]$ and $g \in C_c(\mathbb{R}^d)$,

$$\int g(x) f(t, dx) = \int g(\phi_{0,t}(x)) f_0(dx).$$

Remark 3.2. For $g \in C_c(\mathbb{R}^d)$ and $t \in [0, T]$, $(s, x) \mapsto g(\phi_{s,t}(x))$ is the solution of the transport equation (3.1) in $[0, t] \times \mathbb{R}^d$ with terminal value g at time t, and, hence, f is called the duality solution of (3.2). Equivalently, $f(t, \cdot)$ is the pushforward by $\phi_{0,t}$ of the measure f_0. When f_0 is a probability measure, this means that $f(t, \cdot)$ is the law at time t of the stochastic process $\phi_{0,t}(X_0)$, where X_0 is a random variable with law f_0.

Remark 3.3. The notion of duality solution can be equivalently formulated in relation to nonconservative equations with a right-hand side[7], that is, for $g \in L^1([0, T], C(\mathbb{R}^d))$,

$$-\partial_t u + b(t, x) \cdot \nabla u = g(t, x) \quad \text{in } (0, T) \times \mathbb{R}^d. \tag{3.15}$$

With this perspective, although the object $\text{div}(bf)$ does not make sense as a classical distribution, the equation can still be applied to particular singular test functions, namely, solutions of equations like (3.15). Then the pairing

$$\int_{\mathbb{R}^d} u(T, x) f(T, dx) - \int_{\mathbb{R}^d} u(0, x) f_0(dx)$$
$$+ \int_0^T \int_{\mathbb{R}^d} \underbrace{[-\partial_t u(t, x) + b(t, x) \cdot \nabla u(t, x)]}_{=g(t,x)} f(t, dx) = 0 \tag{3.16}$$

has a sense, because the singular terms collapse into a continuous function, which may be paired with $f(t, \cdot)$.

---

[7] The theory of viscosity solutions of the terminal value problem for (3.15) can be formulated following the theory of the previous subsection with little change.

Remark 3.4. When $d = 1$, the theory for (3.2) can be connected to that for the nonconservative equation (4.2) in Sect. 4 below, in that (4.2) is (up to a time change of b) a primitive of (3.2). Using this relationship, Bouchut and James [22, Theorem 4.3.4] are able to give meaning to the distributional product bf (or $b\partial_x u$ in the next section). More precisely, it is shown that the duality solution in either setting has the following reformulation: there exists $\hat{b} : [0, T] \times \mathbb{R} \to \mathbb{R}$ such that $\hat{b} = b$ a.e., and f (resp. u) is a distributional solution of (3.2) (resp. (4.2)) with b replaced by $\hat{b}$. Extending this concept to multiple dimensions, in which case the direct relationship between (3.2) and (4.2) is not present, seems to be rather difficult.

Theorem 3.8. *There exists a unique duality solution f of (3.2). If, for $\varepsilon > 0$, f^ε is the solution corresponding to b^ε as in (2.3), then, as $\varepsilon \to 0$, f^ε converges weakly in the sense of measures to f. If $1 \le p < \infty$, $f_0, g_0 \in \mathcal{P}_p$, and f and g are the corresponding duality solutions, then, for some $C > 0$ depending on p and the constants in (2.1), $\mathcal{W}_p(f_t, g_t) \le C\mathcal{W}_p(f_0, g_0)$.*

Proof. The existence and uniqueness of duality solutions is a direct consequence of the definition. Moreover, the duality solution identity implies that, for any $R > 0$ and for some $C > 0$ depending on the constants in (2.1), $\|f(t, \cdot)\|_{TV(B_R)} \le \|f_0\|_{TV(B_{R+C})}$. For $0 \le s < t \le T$ and $g \in C_c(\mathbb{R}^d)$, we apply the duality formula with the test function $g \circ \phi_{s,t}$ and obtain the identity

$$\int_{\mathbb{R}^d} g(x) f(t, dx) = \int_{\mathbb{R}^d} g(\phi_{s,t}(x)) f(s, dx).$$

Then, by Lemma 2.2, for some modulus of continuity ω depending on the modulus of continuity for g,

$$\left| \int_{\mathbb{R}^d} g(x) \left[f(t, dx) - f(s, dx) \right] \right| \le \omega(|t - s|) \|f_0\|_{B_{\mathrm{supp}\, g + C}},$$

and we conclude that $f \in C([0, T], \mathcal{M}_{\mathrm{loc,w}})$.

For $R > 0$, define $f_{0,R} := f_0 \mathbf{1}_{B_R}$, and denote by f_R and f_R^ε the duality solutions of (3.2) with respectively b and b^ε and initial condition $f_{0,R}$. It then suffices to prove that, for fixed $R > 0$ as $\varepsilon \to 0$, $f_R^\varepsilon \rightharpoonup f_R$ in the sense of measures. Then, in view of Lemma 2.2, for any $t \in [0, T]$ and $g \in C_c(\mathbb{R}^d)$ for sufficiently large support,

$$\int_{\mathbb{R}^d} g(x) f_R(t, dx) = \int_{B_R} g(\phi_{0,t}(x)) f_0(dx) = \int_{\mathbb{R}^d} g(\phi_{0,t}(x)) f_0(dx)$$
$$= \int_{\mathbb{R}^d} g(x) f(t, dx),$$

and similarly for f^ε.

Let then $g \in C_c(\mathbb{R}^d)$ and $t \in (0, T]$ be fixed, and assume without loss of generality that f_0 has compact support in B_R for some $R > 0$. Then, for $\varepsilon > 0$,

$$\int_{\mathbb{R}^d} g(x) f^\varepsilon(t, dx) = \int g(\phi_{0,t}^\varepsilon(x)) f_0(dx).$$

so that $\|f^\varepsilon\|_{TV} \leq \|f_0\|_{TV}$. Moreover, if $\operatorname{supp} g \subset \mathbb{R}^d \setminus B_{R+C}$ for some $C > 0$ sufficiently large and independent of $\varepsilon > 0$, again by Lemma 2.2,

$$\int_{\mathbb{R}^d} g(x) f^\varepsilon(t, dx) = 0.$$

We may then take a weakly convergent subsequence of f^ε, with limit point $F \in L^\infty([0, T], \mathcal{M})$, and, sending $\varepsilon \to 0$, we obtain that F satisfies the duality solution identity, and therefore $F = f$.

Choose $h_1, h_2 \in C_c(\mathbb{R}^d)$ such that, for all $x, y \in \mathbb{R}^d$, $h_1(x) + h_2(y) \leq |x - y|^p$. Then, if γ is any coupling between f_0 and g_0, we compute, using the duality identity and Lemma 2.2,

$$\int h_1(x) f(t, dx) + \int h_2(y) g(t, dy) = \iint \left(h_1(\phi_{0,t}(x)) + h_2(\phi_{0,t}(y)) \right) \gamma(dx, dy)$$

$$\leq C \iint |x - y|^p \gamma(dx, dy).$$

Taking the infimum over such γ and supremum over such h_1, h_2, and using the dual formulation of the p-Wasserstein distance, we arrive at the estimate for the Wasserstein distances. $\qquad\square$

Remark 3.5. The final estimate can also be proved using the characterization of f and g as laws of certain stochastic processes (see Remark 3.2) and the characterization of the Wasserstein metric in terms of random variables.

We may repeat the above analysis for the second-order conservative equation (3.4), the only difference being the lack of a finite speed of propagation. Therefore, all measures are taken to have finite mass over $\mathbb{R}^d$. Below, $\Phi_{t,0}$ is the stochastic flow satisfying (3.7).

Definition 3.3. A map $f \in C([0, T], \mathcal{M}_w)$ is called a solution of (3.4) if, for all $t \in [0, T]$ and $g \in C_b(\mathbb{R}^d)$,

$$\int g(x) f(t, dx) = \int \mathbb{E}[g(\Phi_{t,0}(x))] f_0(dx).$$

Remark 3.6. Once again, such solutions are called duality solutions because $\mathbb{E}[g \circ \Phi_{t,0}]$ is the solution of (3.3) with terminal value g at time t. If f_0 is a probability measure, then $f(t, \cdot)$ is the law of the stochastic process $\Phi_{t,0}(X_0)$, where X_0 is a random variable with law f_0, independent of the Wiener process W.

The following may be proved exactly as for Theorem 3.8, now invoking the properties of the stochastic flow described by Lemma 2.5.

Theorem 3.9. *There exists a unique duality solution f of (3.4). If, for $\varepsilon > 0$, f^ε is the solution corresponding to b^ε as in (2.3), then, as $\varepsilon \to 0$, f^ε converges weakly in the sense of measures to f. If $1 \leq p \leq \infty$, $f_0, g_0 \in \mathcal{P}_p$, and f and g are the corresponding duality solutions, then, for some $C > 0$ depending on p and the constants in (2.1), $\mathcal{W}_p(f_t, g_t) \leq C \mathcal{W}_p(f_0, g_0)$.*

3.2.2. On the Failure of Renormalization   In view of the formula (3.6), it is immediate that (viscosity) solutions of (3.1) satisfy the renormalization property, that is, if u is a viscosity solution and $\beta : \mathbb{R} \to \mathbb{R}$ is smooth, then $\beta \circ u$ is also a solution. This is related to the existence and uniqueness of the Lipschitz backward flow; indeed, note that, coordinate by coordinate, $\phi_{t,T}(x)$ is the unique viscosity solution of (3.1) with terminal value x at time T.

We contrast this with the renormalization property for the forward, conservative problem (3.2). If b is smooth, then classical computations show that f is a solution if and only if $|f|$, f_+, and f_- are all solutions. Because $f(t, \cdot)$ is the pushforward by f_0 of the flow $\phi_{0,t}$, this can be viewed as a generalized form of injectivity for the flow. For b satisfying (2.1), the backward flow is not guaranteed to be injective, and may in fact concentrate at null sets. We therefore cannot expect renormalization to hold in general.

As a concrete example, take again $b(x) = \operatorname{sgn} x$ on $\mathbb{R}$, and $f_0 = \frac{1}{2}\delta_1 - \frac{1}{2}\delta_{-1}$. Then, for $t > 0$, $f(t, \cdot) = \frac{1}{2}\delta_{(1-t)_+} - \frac{1}{2}\delta_{-(1-t)_+}$, which means that $f(t, \cdot) \equiv 0$ for $t \geq 1$. However, the solution F of (3.2) with $F_0 = |f_0| = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ is equal to $F(t, \cdot) = \frac{1}{2}\delta_{(1-t)_+} + \frac{1}{2}\delta_{-(1-t)_+}$, so that $F(t, \cdot) = \delta_0$ for $t \geq 1$. Thus $F_t \neq |f_t|$ for $t \geq 1$; indeed, $|f_t|$ does not even conserve mass.

The failure of renormalization holds even if we impose $f_0 \in L^1 \cap L^\infty$. For such f_0 and for $b(x) = \operatorname{sgn} x$, we have

$$f(t, dx) = [f_0(x + t)\mathbf{1}\{x > 0\} + f_0(x - t)\mathbf{1}\{x < 0\}] \, dx + \left(\int_{[-t,t]} f_0 \right) d\delta_0(x).$$

Therefore, renormalization fails whenever f_0 is nonzero and odd.

We present one more counterexample to renormalization in which $b \in C$ and $f \in L^1$ (as the previous example shows, even if $f_0 \in L^1$, $f(t, \cdot)$ may not be absolutely continuous with respect to Lebesgue measure due to the concentration of the flow). Take $b(t, x) = 2\operatorname{sgn} x|x|^{1/2}$. The backward flow is given by $\phi_{0,t}(x) = \operatorname{sgn} x(|x|^{1/2} - t)_+^2$ for $(t, x) \in [0, T] \times \mathbb{R}$. For $f_0 \in L^1$, the duality solution is given by

$$f(t, dx) = \left(\int_{[-t^2,t^2]} f_0 \right) \delta_0(dx) + f_0 \left(\operatorname{sgn} x(|x|^{1/2} + t)^2 \right) \frac{|x|^{1/2} + t}{|x|^{1/2}} dx.$$

We then take the odd density $f_0(x) = \operatorname{sgn} x|x|^{1/2}\mathbf{1}_{[-1,1]}(x)$, and the duality solution takes values in L^1:

$$f(t, x) = \operatorname{sgn} x \frac{(|x|^{1/2} + t)^2}{|x|^{1/2}}\mathbf{1}_{[-(1-t)_+^2,(1-t)_+^2]}(x). \tag{3.17}$$

On the other hand, $|f|$ is not the duality solution, or even a distributional solution, since mass is not conserved. The unique duality solution with initial density $|f_0(x)| = |x|^{1/2}\mathbf{1}_{[-1,1]}(x)$ in this case is given by

$$F(t, dx) = \frac{4t^3}{3}\delta_0(dx) + \frac{(|x|^{1/2} + t)^2}{|x|^{1/2}}\mathbf{1}_{[-(1-t)_+^2,(1-t)_+^2]}(x)dx.$$

Remark 3.7. One consequence of the commutator lemma of DiPerna and Lions [40, Lemma II.1] is that, if $f \in L^p$ and $b \in W^{1,q}$ with $\frac{1}{p} + \frac{1}{q} \leq 1$, then the renormalization property is satisfied. The previous example therefore indicates that these conditions cannot be weakened in general. Indeed, even though $f_0 \in L^1 \cap L^\infty$, the solution $f(t, \cdot)$ given by (3.17) belongs to L^p only for $p \in [1, 2)$ when $t > 0$, and the same is true for $\partial_x b$.

3.2.3. Equivalence of Duality and Distributional Solutions

We finish this section by studying the setting where bf can be understood as a distribution, and, therefore, distributional solutions of (3.2) can be considered.

Theorem 3.10. *Assume either that b is continuous, or that $f(t, \cdot) \in L^1_{loc}$ for all $t \in [0, T]$. Then f is a distributional solution of (3.2) if and only if f is the unique duality solution.*

Proof. Suppose f is the unique duality solution. Let $(b^\varepsilon)_{\varepsilon > 0}$ be as in (2.3) and let f^ε be the corresponding solution of (3.2). For $\phi \in C^1_c((0, T) \times \mathbb{R}^d)$, integrating by parts yields

$$\iint_{(0,T) \times \mathbb{R}^d} f^\varepsilon(t, x) \left(-\partial_t \phi(t, x) + b^\varepsilon(t, x) \cdot \nabla \phi(t, x) \right) dt dx = 0.$$

In the case that $b \in C$, we may choose regularizations b^ε that converge locally uniformly to b. By Theorem 3.8, as $\varepsilon \to 0$, f^ε converges weakly in the sense of measures to f, and so we may take $\varepsilon \to 0$ above to obtain

$$\iint_{(0,T) \times \mathbb{R}^d} f(t, dx) \left(-\partial_t \phi(t, x) + b(t, x) \cdot \nabla \phi(t, x) \right) dt = 0.$$

Otherwise, if $f \in L^1_{loc}$, it follows that f^ε converges weakly in L^1_{loc}, and therefore the same is true for $b^\varepsilon f^\varepsilon$ by the dominated convergence theorem. We may then take $\varepsilon \to 0$ in this case as well.

Assume now that f is an arbitrary distributional solution. We aim to show the duality equality in Definition 3.2, and, by a density argument, it suffices to do so for $g \in C_c(\mathbb{R}^d) \cap C^{0,1}(\mathbb{R}^d)$. Let ρ_ε be a standard mollifier as before and set $f_\varepsilon = f * \rho_\varepsilon$. Then f_ε satisfies

$$\partial_t f_\varepsilon - \mathrm{div}(b f_\varepsilon) = \mathrm{div}\, r_\varepsilon,$$

where $r_\varepsilon = (bf) * \rho_\varepsilon - b f_\varepsilon$. For $t \in (0, T]$, let u be the unique Lipschitz viscosity solution of the terminal value problem

$$-\partial_s u + b \cdot \nabla u = 0 \quad \text{in } (0, t) \times \mathbb{R}^d, \quad u(t, \cdot) = g.$$

By the theory in Sect. 3.1, $u(s, x) = g(\phi_{s,t}(x))$ and is Lipschitz continuous with compact support. We then compute

$$\partial_s \int f_\varepsilon(s, x) u(s, x) dx = -\int r_\varepsilon(s, x) \cdot \nabla u(s, x) dx,$$

so that

$$\int f_\varepsilon(t, x) g(x) dx - \int (f_0 * \rho_\varepsilon)(x) g(\phi_{0,t}(x)) dx$$

$$= -\int_0^t \int_{\mathbb{R}^d} r_\varepsilon(s, x) \cdot \nabla u(s, x) dx ds.$$

We may then conclude by proving that $r_\varepsilon \xrightarrow{\varepsilon \to 0} 0$ in L^1_{loc}.

If $f \in L^1_{\text{loc}}$, this is immediate because, as $\varepsilon \to 0$, both $(bf) * \rho_\varepsilon$ and bf_ε converge in L^1_{loc} to bf. If $b \in C$, then, as $\varepsilon \to 0$, both $(bf) * \rho_\varepsilon$ and bf_ε converge locally in total variation to bf. It follows that r_ε converges locally in total variation to 0, but, because $r_\varepsilon \in L^1$ for all $\varepsilon > 0$, the convergence in L^1_{loc} is established.   □

Remark 3.8. Even in the context of Theorem 3.10, the renormalization property can fail. Indeed, this is the case for the final example in the previous section, where both $b \in C$ and $f \in L^1$.

4. The Expansive Regime

We continue our analysis of transport and continuity equations with vector fields b satisfying (2.1), and in this section we study the expansive regime. Reversing the sign appearing in front of the velocity field b, the initial value problem for the continuity equation becomes

$$\partial_t f + \text{div}(b(t, x) f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d \quad \text{and} \quad f(0, \cdot) = f_0, \tag{4.1}$$

and the corresponding dual terminal value problem for the non-conservative transport equation is

$$\partial_t u + b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d \quad \text{and} \quad u(T, \cdot) = u_T. \tag{4.2}$$

Equivalently, we are studying the time-reversed versions of (3.1) and (3.2) (in this case, b is replaced with $b(T - t, \cdot)$). As such, the relevant direction of the flow (2.2) changes in this context: whereas in the previous section, the compressive, backward flow gave rise to the dual solution spaces C and $\mathcal{M}$, here, the expansive, forward flow allows to develop a theory for both (4.1) and (4.2) in Lebesgue spaces. This can also be seen from formal a priori L^p estimates for (4.1) and (4.2), which follow immediately from the lower bound on div b.

The regime for these equations matches those studied by Bouchut et al. [25], in which emphasis is placed on the fact that distributional solutions $f \in C([0, T], L^\infty_{\text{w-}\star}(\mathbb{R}^d))$ of (4.1) are not unique in general. Our approach to these equations is similar, in that we use a particular solution of (4.1) to study, by duality, the transport equation (4.2) and the forward ODE flow to (2.2). We extend the results of [25] by identifying a "good" solution (*reversible solution*, in the terminology of [25]) of (4.2) for any $f_0 \in L^p_{\text{loc}}$, where the continuous solution operator on L^p is stable under regularizations in the weak topology of $C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$.

The terminal value problem (4.2) is then understood both in the dual sense and through the lens of renormalization theory. It is this theory that allows, as in [40], to make sense of the forward ODE flow (2.17) as the right-inverse of the backward flow, completing the program initiated in Sect. 2. As a consequence, we then also obtain the uniqueness of nonnegative distributional solutions of (4.1), and, by extension, a characterization of the reversible solution of [25].

We finish the section by making some remarks about the second-order analogues of (4.1) and (4.2). Unlike in the previous section, we do not have a full solution theory for general second-order equations, unless the ellipticity matrix is uniformly positive (the case which has already been covered by Figalli in [41]) or is degenerate but independent of the space variable.

4.1. The Conservative Equation

The starting point for the study of the conservative equation (4.1) is that solutions in the sense of distributions are not unique (see also [22], [25, Section 6]). We revisit the example, when $d = 1$, $b(t, x) = \text{sgn } x$. Then $f(t, x) := \text{sgn } x \mathbf{1}_{|x| \leq t}$ is a nontrivial distributional solution of (4.1) belonging to $L^1 \cap L^\infty$ with $f(0, \cdot) = 0$. The uniqueness can be seen as a consequence of the compressive nature of the backward flow (2.2), which allows for positive and negative mass to be "cancelled" at time 0, only to appear immediately for $t > 0$. The same phenomenon is what leads to the failure of renormalization for the compressive regime for the continuity equation in Sect. 3.2. In either case, we remark that this particular b belongs to $BV(\mathbb{R})$, while $\partial_x b$ is not absolutely continuous with respect to Lebesgue measure, and so the condition in the work of Ambrosio [5] that div $b \in L^1_{\text{loc}}$ cannot indeed be weakened in general, if one is to hope for renormalization or uniqueness for the continuity equation.

One strategy is to define solutions of (4.1) by duality with the transport equation (3.1) from the compressive setting. With the theory of Sect. 3, for $g \in C^{0,1}_c(\mathbb{R}^d)$, we may define a Lipschitz viscosity solution of the initial value problem

$$\partial_t v + b(t, x) \cdot \nabla v = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad v(0, \cdot) = g$$

(because $\tilde{v}(t, x) := v(T - t, x)$ solves the corresponding terminal value problem (3.1) with velocity $\tilde{b}(t, x) = b(T - t, x)$), and then, formally, for $t > 0$, $\int f(t, x) v(t, x) dx = \int f_0(x) g(x) dx$.

The main problem with this approach is that duality does not define unique solutions, again due to the concentration effect of the backward flow. Taking once more $b(t, x) = \text{sgn } x$, we have, by (3.6),

$$v(t, x) = \begin{cases} g(x - (\text{sgn } x)t), & |x| \geq t, \\ g(0), & |x| \leq t. \end{cases}$$

Therefore, the duality equality fails to give sufficient information to identify f in the cone $\{|x| \leq t\}$, in which v is always constant, regardless of the initial data g. Indeed, the two distributional solutions $f \equiv 0$ and $f(t, x) = \text{sgn } x \mathbf{1}\{|x| \leq t\}$ differ in exactly this cone, in which the Jacobian of the backward flow vanishes.

It is exactly this observation that lead to the notion of "exceptional" solutions of (3.1) and the exceptional set in [25].

Inspired by the work of [25], we instead identify a "good" solution operator acting on all $f_0 \in L_{loc}^p$, $1 \leq p \leq \infty$, by extending the solution formula in the smooth case, which depends on the backward flow studied in Sect. 2, as well as the corresponding Jacobian. In particular, the "good solution" is distinguished by vanishing whenever the Jacobian does. Though our notion of solution turns out to be equivalent to the reversible solutions, our approach differs slightly from that of [25], who work with a general class of "transport flows" that generalize the backward ODE flow. One advantage of our analysis is that we can directly appeal to the various topological properties of the backward flow proved in Sect. 2. Let us also draw analogy with the approach of Crippa and De Lellis [36], in which the regular Lagrangian flow and its properties are studied directly, leading to information about the associated PDEs, although in our setting, we still require PDE techniques to glean further properties of the ODE flow.

4.1.1. Representation Formula   If b is Lipschitz, then the solution of (4.1) is given by

$$f(t, x) = f_0(\phi_{0,t}(x)) J_{0,t}(x), \qquad (4.3)$$

where $\phi_{0,t}(x)$ is the reverse flow defined in Sect. 2 and $J_{0,t}(x) = \det(\nabla_x \phi_{0,t}(x))$ is the corresponding Jacobian. One way to derive this formula is through the Feynman–Kac formula for the reversed time equation

$$-\partial_t \tilde{f} + b(T - t, x) \cdot \nabla \tilde{f} + \operatorname{div}_x b(T - t, x) \tilde{f} = 0 \quad \text{in } (0, T) \times \mathbb{R}^d,$$
$$\tilde{f}(T, \cdot) = f_0,$$

which gives

$$f(t, x) = \tilde{f}(T - t, x) = f_0(\phi_{0,t}(x)) \exp\left(-\int_0^t \operatorname{div} b(s, \phi_{0,s}(x) ds\right), \quad (4.4)$$

and then $J_{0,t}(x) = \exp\left(-\int_0^s \operatorname{div} b(s, \phi_{0,s}(x) ds\right)$.

In the general case where b satisfies (2.1), the formula (4.3) makes sense for arbitrary $f_0 \in L_{loc}^p$, $1 \leq p \leq \infty$. We may then use the various results in Sect. 2 to analyze the stability properties of the solution operator defined by the formula (3.6). We remark in particular that the stability results of Lemma 2.3 depend on the determinant structure of the Jacobian, which is somewhat disguised by the exponential expression in (4.4).

Theorem 4.1. *Let $1 \leq p \leq \infty$, assume that $f_0 \in L_{loc}^p(\mathbb{R}^d)$, and define f by (4.3). Then f is a distributional solution of (4.1). If $1 \leq p < \infty$, $f \in C([0, T], L^p(\mathbb{R}^d))$, and if $p = \infty$, $f \in C([0, T], L_{w\text{-}\star}^\infty(\mathbb{R}^d))$. There exists a constant $C > 0$ depending only on the assumptions in (2.1) such that, for all $R > 0$,*

$$\|f(t, \cdot)\|_{L^p(B_R)} \leq C \|f_0\|_{L^p(B_{R+C})}. \qquad (4.5)$$

If $(b_\varepsilon)_{\varepsilon > 0}$ are as in (2.3) and $(f_\varepsilon)_{\varepsilon > 0}$ are the corresponding solutions of (4.1), then, as $\varepsilon \to 0$, f_ε converges to f weakly in $C([0, T], L_{loc}^p(\mathbb{R}^d))$ if $1 \leq p < \infty$, and weak-$\star$ in L^∞ if $p = \infty$.

Proof. When $p = \infty$, the bound (4.5) follows from the L^∞ bounds for the flow and Jacobian in Lemmas 2.2 and 2.3. We prove the bound when $p < \infty$ for the solutions f_ε of the equation with b_ε as in (2.3), for a constant independent of ε, and then the estimate for f follows after proving the weak convergence result.

For a constant $C > 0$ independent of ε, by Lemmas 2.2 and 2.3, we have $|J_{0,t}^\varepsilon| \leq C$ and $|\phi_{0,t}^\varepsilon(x)| \leq R + C$ for $|x| \leq R$. Then

$$
\begin{aligned}
\int_{B_R} |f^\varepsilon(t,x)|^p dx &= \int_{\mathbb{R}^d} |f_0(\phi_{0,t}^\varepsilon(x))|^p J_{0,t}^\varepsilon(x)^p dx \\
&\leq \|J_{0,t}\|_\infty^{p-1} \int_{B_R} |f_0(\phi_{0,t}^\varepsilon(x))|^p J_{0,t}^\varepsilon(x) dx \\
&\leq C \int_{B_{R+C}} |f_0(x)|^p dx.
\end{aligned}
$$

It suffices to prove the weak convergence of f^ε when $p < \infty$ for $f_0 \in C_c$. In the general case, if $\tilde{f}_0$ is continuous with compact support and we let $\tilde{f}^\varepsilon$ be the solution with b^ε and $\tilde{f}_0$, we have

$$
\left\| f^\varepsilon - \tilde{f}^\varepsilon \right\|_{C([0,T], L^p(B_R))} \leq C \left\| f_0 - \tilde{f}_0 \right\|_{L^p(B_{R+C})},
$$

and we may then choose $\tilde{f}_0$ arbitrarily close to f_0 in L^p_{loc}.

By Lemma 2.2, as $\varepsilon \to 0$, $\phi^\varepsilon \to \phi$ uniformly in $[0,T] \times \mathbb{R}^d$, and therefore $f_0 \circ \phi_{0,t}^\varepsilon)$ converges uniformly to $f_0 \circ \phi_{0,t}$. In view of Lemma 2.3, f^ε converges weakly in the sense of distributions (and therefore, in the sense of locally bounded Borel measures) to f. Since f^ε is bounded in $L^\infty([0,T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$, the convergence is actually weak in $L^\infty([0,T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$.

If $p = \infty$, then, in particular, $f^\varepsilon \in C([0,T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$ for $p < \infty$, uniformly in ε, and we have the convergence as $\varepsilon \to 0$ in the sense of distributions to f. In this case, $f \in L^\infty_{\mathrm{loc}}([0,T] \times \mathbb{R}^d)$, and so the convergence is weak-$\star$ in L^∞_{loc}.

Given $g \in C_c^1((0,T) \times \mathbb{R}^d)$, integrating by parts gives

$$
\iint_{[0,T] \times \mathbb{R}^d} f^\varepsilon(t,x) \left[\partial_t \phi(t,x) + b^\varepsilon(t,x) \cdot \nabla \phi(t,x) \right] dx dt = 0.
$$

As $\varepsilon \to 0$, the bracketed expression converges a.e. to $\partial_t \phi(t,x) + b(t,x) \cdot \nabla \phi(t,x)$, and so converges weakly in L^q for all $1 \leq q < \infty$ by the dominated convergence theorem. We may therefore send $\varepsilon \to 0$, using the weak convergence of f^ε, to deduce that f is a distributional solution. This implies in particular that $f \in C([0,T], L^p_{\mathrm{w}}(\mathbb{R}^d))$, or $C([0,T], L^\infty_{\mathrm{w}\text{-}\star}(\mathbb{R}^d))$ if $p = \infty$.

To show that $f \in C([0,T], L^p(\mathbb{R}^d))$ when $p < \infty$, we may again consider $f_0 \in C_c(\mathbb{R}^d))$ without loss of generality. Then $f_0 \circ \phi_{0,\cdot} \in C([0,T] \times \mathbb{R}^d)$, while $J_{0,\cdot} \in C([0,T], L^1_{\mathrm{loc}}(\mathbb{R}^d))$ by Lemma 2.3, and the result follows. $\square$

Remark 4.1. In view of the stability results of Theorem 4.1 above, this "good" solution coincides with the notion of reversible solutions in [22,25]. We refer to it in the sequel as the BJM solution.

The following is immediately obtained from the formula (4.3):

Corollary 4.1. *If f is a BJM solution of* (4.1)*, then so is $|f|$.*

Corollary 4.1 is in direct contrast to the continuity equation in the compressive setting of the previous section, where renormalization fails. Its proof depends on the formula for the BJM solution; indeed, despite the weak stability result in Theorem 4.1, this renormalization property cannot be proved by regularization, since we only have the weak convergence as $\varepsilon \to 0$ of f_ε to f. At present, we do not know whether the convergence is strong in L^p. This turns out to be equivalent to the strong convergence in L^1_{loc} of the Jacobians, and therefore, in view of Proposition 2.1, we have the following when $d = 1$.

Theorem 4.2. *Assume $d = 1$, $f_0 \in L^p_{\mathrm{loc}}(\mathbb{R})$ for $p < \infty$, $(b^\varepsilon)_{\varepsilon>0}$ is as in* (2.3)*, and f^ε is the corresponding solution of* (4.1)*. Then, as $\varepsilon \to 0$, f^ε converges strongly in $C([0, T], L^p_{\mathrm{loc}}(\mathbb{R}))$ to f.*

Proof. Just as in the proof of Theorem 4.1, we may assume without loss of generality that $f_0 \in C_c(\mathbb{R})$. In that case, f^ε is bounded in L^1 and L^∞, and so the strong L^p convergence reduces to the strong convergence of $J^\varepsilon_{0,\cdot}$ to $J_{0,\cdot}$ in $L^1_{\mathrm{loc}}([0, T] \times \mathbb{R})$ from Proposition 2.1. $\qquad\square$

4.1.2. Vanishing Viscosity Approximation   The BJM solution above also arises from vanishing viscosity limits, that is, the limit as $\varepsilon \to 0$ of solutions of

$$\partial_t f^\varepsilon - \frac{\varepsilon^2}{2}\Delta f^\varepsilon + \mathrm{div}(b(t, x) f^\varepsilon) = 0 \quad \text{in } [0, T] \times \mathbb{R}^d \quad \text{and} \quad f^\varepsilon(0, \cdot) = f_0, \quad (4.6)$$

which has as its unique solution

$$f^\varepsilon(t, x) := \mathbb{E}[f_0(\phi^\varepsilon_{0,t}(x)) J^\varepsilon_{0,t}(x)], \quad (4.7)$$

where now ϕ^ε and J^ε denote respectively the stochastic flow and Jacobian from (2.23), corresponding to Proposition 2.3.

The proof of the following result follows from Proposition 2.3, and is proved almost exactly as for Theorem 4.1.

Theorem 4.3. *The function f^ε defined by* (4.6) *belongs to $C([0, T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$ if $1 \leq p < \infty$ and $C([0, T], L^\infty_{\mathrm{w}\text{-}\star}(\mathbb{R}^d))$ if $p = \infty$, and, as $\varepsilon \to 0$, f^ε converges weakly in those spaces to f.*

4.2. The Nonconservative Equation

The next step is the study of the terminal value problem (4.2). Unlike the transport equation (3.1) with velocity $-b$, which was solved in the space of continuous functions, we cannot define L^p solutions in the distributional sense, as the product $b \cdot \nabla u = \mathrm{div}(bu) - (\mathrm{div}\, b)u$ does not make sense when $\mathrm{div}\, b$ is merely a measure. Instead, we initially characterize solutions by duality with (4.1), which can be seen as a way of restricting the class of test functions to deal with the singularities in b (see Remark 3.3).

4.2.1. L^p and BV Estimates   We will first prove a priori L^p and BV estimates for the solution of (4.2), assuming all the data and solutions are smooth. The BV estimates in particular are crucial to establishing the *strong* convergence in L^p of regularized solutions to a unique limit, which will be the duality solution, adjoint to the equation (4.1). The BV estimate appears already in [25, Lemma 4.4]. We present an alternate proof here, which is similar to the one for second-order equations we prove later.

Lemma 4.1. *Assume b is smooth and satisfies* (2.1)*, and let u be a smooth solution of* (4.2)*. Then, for all $1 \leq p \leq \infty$, there exist $C = C_{p,R} \in L^1_+([0, T])$ and $C_R > 0$ depending only on the bounds in* (2.1) *such that, for all $0 \leq t \leq T$,*

$$\|u(t, \cdot)\|_{L^p(B_R)} \leq \exp\left(\int_0^t C(s)ds\right) \|u_T\|_{L^p(B_{R+C})}$$

and

$$\|u(t, \cdot)\|_{BV(B_R)} \leq \exp\left(\int_0^t C(s)ds\right) \|u_T\|_{BV(B_{R+C})}.$$

Proof. We assume that u_T has compact support, and, therefore, in view of the finite speed of propagation property, so does u. The general result for L^p_{loc} and BV_{loc} is proved similarly.

The L^∞ bound is a consequence of the maximum principle. For $p < \infty$, we compute

$$\frac{\partial}{\partial t} \int_{\mathbb{R}^d} |u(t, x)|^p dx = \int_{\mathbb{R}^d} \text{div } b(t, x)|u(t, x)|^p dx \geq -C_0(t)d \int_{\mathbb{R}^d} |u(t, x)|^p dx,$$

and the L^p bound follows from Grönwall's inequality.

Now, for $t \leq T$ and $x, z \in \mathbb{R}^d$, set $w(t, x, z) = \nabla u(t, x) \cdot z$. Then w satisfies

$$-\partial_t w + b \cdot \nabla_x w + (z \cdot \nabla)b \cdot \nabla_z w = 0.$$

Since b and w are smooth, the renormalization property holds for this transport equation, and so a simple regularization argument shows, in the sense of distributions,

$$\partial_t |w| + b \cdot \nabla_x |w| + (z \cdot \nabla)b \cdot \nabla_z |w| = 0.$$

Define $\phi(z) = e^{-|z|^2}$. Then

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \phi(z)b(t, x) \cdot \nabla_x |w(t, x, z)| dx dz$$
$$= -\iint_{\mathbb{R}^d \times \mathbb{R}^d} \phi(z)\text{div } b(t, x)|w(t, x, z)| dx dz$$

and

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \phi(z)(z \cdot \nabla)b(t, x) \cdot \nabla_z |w(t, x, z)| dx dz$$
$$= -\iint_{\mathbb{R}^d \times \mathbb{R}^d} [\nabla\phi(z) \cdot (z \cdot \nabla b(t, x)) + \phi(z)\text{div } b(t, x)] |w(t, x, z)| dx dz.$$

Therefore, by Lemma 2.1,

$$
\partial_t \iint_{\mathbb{R}^d \times \mathbb{R}^d} |w(t, x, z)| \phi(z) dx dz
$$

$$
= \iint_{\mathbb{R}^d \times \mathbb{R}^d} [\nabla \phi(z) \cdot (z \cdot \nabla b(t, x)) + 2\phi(z)\mathrm{div}\, b(t, x)] |w(t, x, z)| dx dz
$$

$$
= \iint_{\mathbb{R}^d \times \mathbb{R}^d} 2e^{-|z|^2} [\mathrm{div}\, b(t, x) - \nabla b(t, x)z \cdot z] dx dz
$$

$$
\geq -2(d-1)C_0(t) \iint_{\mathbb{R}^d \times \mathbb{R}^d} e^{-|z|^2} |w(t, x, z)| dx dz.
$$

The result follows from Grönwall's lemma and the fact that

$$
\iint_{\mathbb{R}^d \times \mathbb{R}^d} e^{-|z|^2} |w(t, x, z)| dx dz = c_0 \int_{\mathbb{R}^d} |\nabla u(t, x)| dx,
$$

where the constant $c_0 = \int_{\mathbb{R}^d} e^{-|z|^2} |v \cdot z| dz$ is independent of the choice of $|v| = 1$ by rotational invariance. $\square$

4.2.2. Duality Solutions   Proceeding by duality with the conservative forward equation, and using the BV-estimates above, then gives the following.

Theorem 4.4. *Assume $1 \leq p \leq \infty$ and $u_T \in L^p_{\mathrm{loc}}$. Then there exists a unique function $u \in C([0, T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$ (or in $C([0, T], L^\infty_{\mathrm{w}\text{-}\star}(\mathbb{R}^d))$ if $p = \infty$) such that, if $(b^\varepsilon)_{\varepsilon>0}$ is as in (2.3) and u^ε denotes the corresponding solution of (4.2), then, as $\varepsilon \to 0$, u^ε converges strongly in $C([0, T], L^p(\mathbb{R}^d))$ for $p < \infty$ and weak-$\star$ in L^∞ to u. Moreover, the solution map $u_T \mapsto u$ is linear, order-preserving, and continuous in $L^p_{\mathrm{loc}}(\mathbb{R}^d)$). If $s \in [0, T)$, $f_s \in L^{p'}(\mathbb{R}^d)$ and $f \in C([s, T], L^{p'}(\mathbb{R}^d))$ (or $C([s, T], L^\infty_{\mathrm{w}\text{-}\star}(\mathbb{R}^d))$ if $p = 1$) is the BJM solution of (4.1) with initial data $f(s, \cdot) = f_s$, then*

$$
\int_{\mathbb{R}^d} u(s, x) f_s(x) dx = \int_{\mathbb{R}^d} u_T(x) f(T, x) dx.
$$

Remark 4.2. The function u corresponds with the notion of duality solution presented in [25] whenever u_T (and therefore $u(t, \cdot)$ for $t < T$) belongs to BV_{loc}.

Proof. By Lemma 4.1, $(u^\varepsilon)_{\varepsilon>0}$ is bounded uniformly in $C([0, T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$, and so, along a subsequence, converges weakly as $\varepsilon \to 0$ to some u satisfying the same bounds.

In order to see that the convergence is strong, note that it suffices, by the L^p-boundedness of solution operator implied by Lemma 4.1, to assume that $u_T \in C_c(\mathbb{R}^d)$. We then have u^ε bounded in $L^\infty([0, T], BV(\mathbb{R}^d))$ independently of ε. The identity $\partial_t u^\varepsilon = -b^\varepsilon \cdot \nabla u^\varepsilon$ then implies that, for any $t_1 < t_2 \leq T$ and $R > 0$,

$$
\left\| u^\varepsilon(t_1, \cdot) - u^\varepsilon(t_2, \cdot) \right\|_{L^1(B_R)} \leq \|b\|_{L^\infty(B_R)} \sup_{t \in [0, T]} \left\| \nabla u^\varepsilon \right\|_{L^1(B_R)} |t_1 - t_2|.
$$

This, along with the uniform BV estimates, implies that $(u^\varepsilon)_{\varepsilon>0}$ is precompact in $C([0, T], L^1_{\mathrm{loc}}(\mathbb{R}^d))$, and, because of the uniform L^∞-bound, precompact in

$C([0, T], L_{\text{loc}}^p(\mathbb{R}^d))$ for any $p \in [1, \infty)$. It therefore follows that any weakly convergent subsequence actually converges strongly.

If f^ε is the solution of (4.1) with $f^\varepsilon(s, \cdot) = f_s$, then classical computations involving integration by parts give

$$\int_{\mathbb{R}^d} u^\varepsilon(s, x) f_s(x) dx = \int_{\mathbb{R}^d} u_T(x) f^\varepsilon(T, x) dx.$$

Sending $\varepsilon \to 0$ along a subsequence and using the weak convergence of f^ε and strong convergence of u^ε shows that any limit point u must satisfy the duality identity with f, and is therefore unique. We conclude that the full sequence converges strongly. As before, when $p = \infty$, we obtain the same result since then also $u \in C([0, T], L_{\text{loc}}^p(\mathbb{R}^d))$ for any $p < \infty$. $\square$

Remark 4.3. If $u_T \in BV_{\text{loc}}$, then the duality solution u of (4.2) satisfies $\nabla u \in L^\infty([0, T], \mathcal{M}_{\text{loc}}(\mathbb{R}^d))$. Note, however, that this is still not enough to make sense of u as a distributional solution, unless b is continuous; see also Remark 3.4.

4.2.3. Renormalization   In Sect. 3, the renormalization property for solutions of the transport equation (3.1) followed from the formula (3.6). We prove a similar renormalization property for the transport equation (4.2) in the expansive regime. Here, it depends on the strong convergence in L^p of regularizations.

Theorem 4.5. *Let* $1 \le p \le \infty$ *and* $u_T \in L_{\text{loc}}^p(\mathbb{R}^d)$, *and let* $u \in C([0, T], L_{\text{loc}}^p(\mathbb{R}^d))$ *be the duality solution of* (4.2). *Assume* $\beta : \mathbb{R} \to \mathbb{R}$ *is smooth and satisfies* $|\beta(r)| \le C(1 + |r|^\alpha)$ *for some* $C, \alpha > 0$. *Then* $\beta \circ u = C([0, T], L_{\text{loc}}^{p/\alpha}(\mathbb{R}^d))$ *is the duality solution of* (4.2) *with terminal value* $\beta(u(T, \cdot)) = \beta \circ u_T$.

Proof. The proof is an easy consequence of regularization of b as in (2.3), and the passage to the limit follows from the strong convergence of u^ε to u. $\square$

4.3. The Forward ODE Flow

We finally return to the study of the flow (2.2), in particular for the forward direction. A candidate for the object $\phi_{t,s}(x)$, $t > s$, a.e. x was already identified in Proposition 2.2 as the right inverse of the backward flow—note that the full measure set of $x \in \mathbb{R}^d$ depends on s and t. We now connect this right-inverse with the transport equation (4.2), and exploit the renormalization property to identify $\phi_{t,s}(x)$ as a regular Lagrangian flow, that is, for a.e. $x \in \mathbb{R}^d$, an absolutely continuous solution of the integral equation for (2.2) with control on the compressibility.

4.3.1. Properties of the Right Inverse   We first record more properties of the right-inverse of the backward flow identified in Proposition 2.2. From now on, for $0 \le s \le t \le T$, we always denote by $\phi_{t,s}$ the version of the right-inverse of $\phi_{s,t}$ which is continuous almost everywhere (such a version is guaranteed to exist by Proposition 2.2).

Theorem 4.6. *For any* $t \in (0, T]$, $[0, t] \times \mathbb{R}^d \ni (s, x) \mapsto \phi_{t,s}(x)$ *is (coordinate-by-coordinate) the duality solution of* (4.2) *with terminal value x at time t. For all* $1 \le p < \infty$,

$$\phi_{\cdot,s} \in C([s, T], L^p_{\text{loc}}(\mathbb{R}^d)) \quad and \quad \phi_{t,\cdot} \in C([0, t], L^p_{\text{loc}}(\mathbb{R}^d)),$$

and there exists a constant $C > 0$ *such that, for all* $0 \le s \le t \le T$ *and* $x \in \mathbb{R}^d$,

$$|\phi_{t,s}(x)| \le C(1 + |x|) \quad and \quad \left\| \phi_{t,s} \right\|_{BV_{\text{loc}}} \le C.$$

Finally, if $(b^\varepsilon)_{\varepsilon > 0}$ *is as in* (2.3) *and* $\phi^\varepsilon_{t,s}$ *is the corresponding forward flow, then, for all* $1 \le p < \infty$,

$$\lim_{\varepsilon \to 0} \phi^\varepsilon_{\cdot,s} = \phi_{\cdot,s} \quad strongly\ in\ C([s, T], L^p_{\text{loc}}(\mathbb{R}^d))$$

and

$$\lim_{\varepsilon \to 0} \phi^\varepsilon_{t,\cdot} = \phi_{t,\cdot} \quad strongly\ in\ C([0, t], L^p_{\text{loc}}(\mathbb{R}^d)),$$

and the convergence also holds in the weak-⋆ sense in L^∞_{loc}.

Proof. For $\varepsilon > 0$ and $(b^\varepsilon)_{\varepsilon > 0}$ as in (2.3), it is standard that, for $t \in (0, T]$, the vector-valued solution of

$$\frac{\partial u^\varepsilon}{\partial s} + b^\varepsilon \cdot \nabla u^\varepsilon = 0 \quad \text{in } (0, t) \times \mathbb{R}^d, \quad u^\varepsilon(t, x) = x$$

is given by $u^\varepsilon(s, x) = \phi^\varepsilon_{t,s}(x)$ for $s \in [0, t]$, where ϕ^ε is the flow corresponding to b^ε. By Theorem 4.4, we have the given convergence statements, as $\varepsilon \to 0$, of ϕ^ε to the vector valued duality solution u of (4.2) in $[0, t] \times \mathbb{R}^d$ with terminal value $u(t, \cdot) = x$.

The flow property for smooth b^ε yields, for $0 \le s \le t \le T$ and $x \in \mathbb{R}^d$, $\phi^\varepsilon_{s,t}(\phi^\varepsilon_{t,s}(x))$. By Lemma 2.2 and the above strong L^p-convergence statement, we may take $\varepsilon \to 0$ to obtain $\phi_{s,t}(u(s, x)) = x$, and then, by Proposition 2.2, we must have $u(s, x) = \phi_{t,s}(x)$. The other statements now follow immediately in view of Theorem 4.4. Note that we are using that, for $s \in [0, T)$, the map $[s, T] \times \mathbb{R}^d \ni (t, x) \mapsto \phi_{t,s}(x)$ is the duality solution of the initial value problem

$$\frac{\partial \tilde{u}}{\partial t} - b(t, x) \cdot \nabla \tilde{u} = 0 \quad \text{in } [s, T] \times \mathbb{R}^d, \quad \tilde{u}(s, x) = x,$$

whose theory can be treated exactly as for (4.2). □

4.3.2. The Regular Lagrange Property   We now observe that there is a representation formula for the duality solution of the transport equation (4.2).

Theorem 4.7. *Let $1 \leq p \leq \infty$. Then there exists a constant $C > 0$ depending only on p and the constant in (2.1) such that, for all $F \in L^p_{\text{loc}} \cap C$, $R > 0$, and $0 \leq s \leq t \leq T$,*

$$\left\| F \circ \phi_{t,s} \right\|_{L^p(B_R)} \leq C \left\| F \right\|_{L^p(B_{R+C})}. \tag{4.8}$$

In particular, for any $A \subset \mathbb{R}^d$ with finite Lebesgue measure,

$$\left| \left\{ x : \phi_{t,s}(x) \in A \right\} \right| \leq C|A|. \tag{4.9}$$

If $u_T \in L^p_{\text{loc}}(\mathbb{R}^d)$, then the duality solution of (4.2) is given by

$$u(t, x) = u_T(\phi_{T,t}(x)). \tag{4.10}$$

If u_T has a version which is continuous almost everywhere, then, for $t < T$, $u(t, \cdot)$ also has a version that is continuous almost everywhere.

Remark 4.4. When u_T is not continuous, then (4.10) must be interpreted as the continuous extension of the operator $u_T \mapsto u_T \circ \phi_{T,t}$ to $u_T \in L^p_{\text{loc}}$, which is well-defined in view of the estimate (4.8).

Remark 4.5. The estimate (4.9) is called the regular Lagrange property. It reinforces the fact that $\phi_{t,s}$ does not concentrate in sets of measure zero.

Remark 4.6. The propagation of almost-everywhere continuity is a consequence of the same property for the forward flow (Proposition 2.2). Note that it is not true in general that a function $u \in BV_{\text{loc}}(\mathbb{R}^d)$ is continuous almost everywhere, unless $d = 1$.

Proof of Theorem 4.7. For continuous u_T, the representation formula is an immediate consequence of the renormalization property Theorem 4.5 and Theorem 4.6. The estimate (4.8) then follows from Theorem 4.4, and (4.9) is obtained by taking $p = 1$ and $F = \mathbf{1}_A$.

For the claim about almost everywhere continuity, define

$$A := \left\{ y \in \mathbb{R}^d : u_T \text{ is not continuous at } y \right\}.$$

Then $|A| = 0$, and then (4.9) gives, for $0 \leq t < T$,

$$\left| \left\{ x \in \mathbb{R}^d : u_T \text{ is not continuous at } \phi_{T,t}(x) \right\} \right| = 0.$$

It follows that u_T is continuous at $\phi_{T,t}(x)$ for a.e. x. By Proposition 2.2, $\phi_{T,t}$ is continuous almost everywhere, and the result follows. □

Recalling the duality relationship between (4.1) and (4.2) from Theorem 4.4, we then have

Corollary 4.2. *For any $1 \leq p \leq \infty$ and $f_0 \in L^p_{\text{loc}}(\mathbb{R}^d)$, the BJM reversible solution f of (4.1) is given at time $t > 0$ by $\phi^{\#}_{t,0} f_0$.*

Remark 4.7. The regular Lagrange property says that the measure $\phi_{t,0}^{\#} f_0$ is well-defined and absolutely continuous with respect to Lebesgue measure, with a density in L_{loc}^{p}. If f_0 is the density for a probability measure, that is, $f_0 \in L_{+}^{1}(\mathbb{R}^d)$ and $\int f_0 = 1$, then $f(t, \cdot)$ is the law at time t of the stochastic process $\phi_{t,0}(X)$, where X is a random variable with density f_0.

A consequence of renormalization and the regular Lagrange property is the fact that the forward flow $\phi_{t,s}$ solves the ODE (2.2) for a.e. initial $x \in \mathbb{R}^d$. A first step is the following lemma.

Lemma 4.2. *For all $p \in [1, \infty)$ and $s \in [0, T)$, $\{(t, x) \mapsto b(t, \phi_{t,s}(x))\} \in L^1([0, T], L_{\text{loc}}^{p}(\mathbb{R}^d))$. If $(b^{\varepsilon})_{\varepsilon > 0}$ is as in (2.3) and $(\phi^{\varepsilon})_{\varepsilon > 0}$ is the corresponding flow, then, for all $R > 0$,*

$$\lim_{\varepsilon \to 0} \int_{s}^{T} \left\| b^{\varepsilon}(t, \phi_{t,s}^{\varepsilon}) - b(t, \phi_{t,s}) \right\|_{L^p(B_R)} dt = 0.$$

Proof. The first claim follows from (4.8): there exists $C > 0$ independent of s and R such that, for all $t \in [0, T]$, $\left\| b(t, \phi_{t,s}) \right\|_{L^p(B_R)} \leq C \left\| b(t, \cdot) \right\|_{L^p(B_{R+C})}$.

For $\delta > 0$ and $0 \leq s \leq t \leq T$, we write

$$\left\| b^{\varepsilon}(t, \phi_{t,s}^{\varepsilon}) - b(t, \phi_{t,s}) \right\|_{L^p(B_R)} \leq \left\| b^{\varepsilon}(t, \phi_{t,s}^{\varepsilon}) - b^{\delta}(t, \phi_{t,s}^{\varepsilon}) \right\|_{L^p(B_R)}$$
$$+ \left\| b^{\delta}(t, \phi_{t,s}^{\varepsilon}) - b^{\delta}(t, \phi_{t,s}) \right\|_{L^p(B_R)} + \left\| b^{\delta}(t, \phi_{t,s}) - b(t, \phi_{t,s}) \right\|_{L^p(B_R)}.$$

By (4.8), for some $C > 0$ independent of δ, ε, s, and t,

$$\left\| b^{\varepsilon}(t, \phi_{t,s}^{\varepsilon}) - b^{\delta}(t, \phi_{t,s}^{\varepsilon}) \right\|_{L^p(B_R)} \leq C \left\| b^{\varepsilon}(t, \cdot) - b^{\delta}(t, \cdot) \right\|_{L^p(B_{R+C})}$$

and

$$\left\| b^{\delta}(t, \phi_{t,s}) - b(t, \phi_{t,s}) \right\|_{L^p(B_R)} \leq C \left\| b^{\delta}(t, \cdot) - b(t, \cdot) \right\|_{L^p(B_{R+C})}.$$

The smoothness of b^{δ} implies that, for all $t \in [s, T]$, as $\varepsilon \to 0$, $b^{\delta}(t, \phi_{t,s}^{\varepsilon})$ converges a.e. to $b^{\delta}(t, \phi_{t,s})$. Sending $\varepsilon \to 0$ and using dominated convergence, we thus have

$$\limsup_{\varepsilon \to 0} \int_{s}^{T} \left\| b^{\varepsilon}(t, \phi_{t,s}^{\varepsilon}) - b(t, \phi_{t,s}) \right\|_{L^p(B_R)} dt$$
$$\leq C \int_{s}^{T} \left\| b^{\delta}(t, \cdot) - b(t, \cdot) \right\|_{L^p(B_{R+C})} dt.$$

The proof of the claim is finished upon sending $\delta \to 0$ and again using dominated convergence. $\square$

Theorem 4.8. *Fix $1 \leq p < \infty$ and $s \in [0, T)$. Then*

$$\{(t, x) \mapsto \phi_{t,s}(x)\} \in L_{\text{loc}}^{p}(\mathbb{R}^d, W^{1,1}([0, T])),$$

and, for a.e. $x \in \mathbb{R}^d$, $[s, T] \ni t \mapsto \phi_{t,s}(x)$ is an absolutely continuous solution of

$$\phi_{t,s}(x) = x + \int_{s}^{t} b(r, \phi_{r,s}(x)) dr.$$

If $(b^\varepsilon)_{\varepsilon>0}$ satisfy (2.3) and ϕ^ε is the corresponding flow, then, for all $R > 0$,

$$\lim_{\varepsilon \to 0} \left\| \phi^\varepsilon_{\cdot,s} - \phi_{\cdot,s} \right\|_{L^p(B_R, W^{1,1}([s,T])} = 0.$$

For all $0 \leq r \leq s \leq t \leq T$, $\phi_{t,r} = \phi_{t,s} \circ \phi_{s,r}$ a.e.

Remark 4.8. The fact that $\partial_t \phi_{t,\cdot} \in L^1$ is due to the fact that we are assuming the weakest possible integrability of b in the time variable. If $b \in L^q$ for some $q > 1$, then the forward flow belongs to $W^{1,p}$ for any $p \leq q$.

Remark 4.9. The composition $\phi_{t,s} \circ \phi_{s,r}$ is made sense of due to (4.8) and the fact that the forward flow takes values in $L^p_{\text{loc}}(\mathbb{R}^d)$.

Proof of Theorem 4.8. For $\varepsilon > 0$, we have $\partial_t \phi^\varepsilon_{t,s}(x) = b^\varepsilon(t, \phi^\varepsilon_{t,s}(x))$. By Lemma 4.2, sending $\varepsilon \to 0$, we see that the distribution $\partial_t \phi_{t,s}(x)$ satisfies, in the distributional sense, $\partial_t \phi_{t,s}(x) = b(t, \phi_{t,s}(x))$, and therefore, for all $R > 0$,

$$\left\| \int_s^T |\phi_{t,s}| dt \right\|_{L^p(B_R)} \leq \int_s^T \|\phi_{t,s}\|_{L^p(B_R)} dt < \infty.$$

The convergence claim and the solvability of the ODE follow immediately in view of the fact that $\phi^\varepsilon_{s,s}(x) = \phi_{s,s}(x) = x$ for all $\varepsilon > 0$ and $x \in \mathbb{R}^d$.

To prove the last claim, we note that the equality $\phi_{r,t} \circ \phi_{t,r} = \text{Id}$ holds as functions in L^p_{loc}, and, in view of the flow property of the backward flow,

$$\phi_{r,t} \circ (\phi_{t,s} \circ \phi_{s,r}) = \phi_{r,s} \circ \phi_{s,t} \circ \phi_{t,s} \circ \phi_{s,r} = \phi_{r,s} \circ \phi_{s,r} = \text{Id}.$$

It follows from Proposition 2.2 that $\phi_{t,r} = \phi_{t,s} \circ \phi_{s,r}$ a.e., as desired. $\quad\square$

We recall that Proposition 2.2 implies that any right-inverse of the backward flow is determined uniquely almost everywhere. We remark here that this property actually follows from the duality between the transport and continuity equations.

Theorem 4.9. *Assume $\psi \in C([0, t], L^p_{\text{loc}}(\mathbb{R}^d))$ satisfies $\phi_{s,t}(\psi_s(x)) = x$ for all $s \in [0, t]$, for a.e. $x \in \mathbb{R}^d$. Then $\psi = \phi_{t,\cdot}$.*

Proof. It suffices to show that $u(t, x) = \psi_s(x)$ is the unique (vector-valued) duality solution of (4.2) with terminal data equal to x at time t.

Fix $g \in C_c(\mathbb{R}^d)$. For a.e. $x \in \mathbb{R}^d$, if $y = \psi_t(x)$, we have $\phi_{s,t}(y) = x$ by assumption. Therefore, the change of variables formula yields

$$\int_{\mathbb{R}^d} g(x)\psi_t(x)dx = \int_{\mathbb{R}^d} g(\phi_{s,t}(y))y J_{s,t}(y)dy = \int_{\mathbb{R}^d} f(t, y)ydy,$$

where f is the BJM solution of the forward continuity equation with initial condition g at time s. $\quad\square$

Remark 4.10. A corresponding result characterizing $\phi_{\cdot,s}$ on $[s, T]$ follows in exactly the same way, by considering the duality between the IVP and TVP for, respectively, an appropriate transport and continuity equation.

Remark 4.11. The uniqueness result above demonstrates that the right-inverse property is a crucial property of the forward flow. In other words, it implies that $\phi_{t,s}$ solves the ODE, that it solves the transport PDE in the duality sense, and that it has the regularity properties laid out in Theorems 4.7 and 4.8.

4.4. Characterizations

We now present alternative ways to characterize the solutions of the forward continuity and backward transport equations identified above. Although the PDE (4.2) does not make sense as a distribution, we nevertheless can characterize solutions in a PDE sense through the use of sup- and inf-convolutions. The propagation of almost-everywhere continuity proved in Theorem 4.7 is a crucial ingredient.

By using this characterization in duality with the conservative equation, we then show that nonnegative distributional solutions of (4.1) are unique, and therefore equal to the solution identified by the formula (4.3). As a consequence, we finally conclude with the uniqueness of regular Lagrangian flows, forward in time, of the ODE (2.2).

4.4.1. The Nonconservative Equation: sup and inf Convolutions   We now identify those regularizations that will lead to a PDE characterization for solutions of the equation (4.2). Given $\delta > 0$ and $u \in L^\infty(\mathbb{R}^d)$, we define the sup- and inf-convolutions

$$u^\delta(x) := \operatorname*{ess\,sup}_{y \in \mathbb{R}^d} \left\{ u(y) - \frac{1}{2\delta}|x - y|^2 \right\}$$

and

$$u_\delta(x) := \operatorname*{ess\,inf}_{y \in \mathbb{R}^d} \left\{ u(y) + \frac{1}{2\delta}|x - y|^2 \right\}.$$

These regularizations are common in the theory of viscosity solutions, or generally for equations satisfying a maximum principle in spaces of continuous functions. The supremum and infimum must be essential, because u is only defined almost everywhere.

Lemma 4.3. *Assume that $u \in L^\infty(\mathbb{R}^d)$ is continuous almost everywhere. Then, for all $\delta > 0$, u_δ, u^δ are globally Lipschitz with constant*

$$(\operatorname{ess\,sup} u - \operatorname{ess\,inf} u)^{1/2} \delta^{-1/2},$$

and

$$u_\delta \leq u \leq u^\delta \quad a.e.$$

As $\delta \to 0$, u^δ decreases to u and u_δ increases to u a.e. Finally, the ess sup and ess inf in the definitions of u^δ and u_δ can be restricted to respectively $y \in B_{R^\delta(x)}(x)$ and $B_{R_\delta(x)}(x)$, where

$$R^\delta(x) = 2(u^{2\delta}(x) - u^\delta(x))^{1/2} \delta^{1/2}$$

and

$$R_\delta(x) = 2(u_\delta(x) - u_{2\delta}(x))^{1/2} \delta^{1/2}.$$

Proof. Fix $x \in \mathbb{R}^d$ and $r > 0$. We thus have

$$u^\delta(x) \geq \operatorname*{ess\,sup}_{y \in B_r(x)} u(y) - \frac{r^2}{2\delta}.$$

Sending $r \to 0$, we see that $u^\delta(x) \geq u(x)$ whenever u is continuous at x, and therefore $u^\delta \geq u$ a.e. Similarly, $u_\delta \leq u$ a.e.

We now observe that, if $R > (\operatorname{ess\,sup} u - \operatorname{ess\,inf} u)^{1/2}$, then, for a.e. $y \notin B_{R\delta^{1/2}}$,

$$u(y) - \frac{|x - y|^2}{2\delta} \leq \operatorname{ess\,sup} u - R^2 < \operatorname{ess\,inf} u \leq u^\delta(x).$$

By also using a similar argument for u_δ, we see that

$$u^\delta(x) := \operatorname*{ess\,sup}_{|y-x| \leq R\delta^{1/2}} \left\{ u(y) - \frac{1}{2\delta}|x - y|^2 \right\}$$

and

$$u_\delta(x) := \operatorname*{ess\,inf}_{|y-x| \leq R\delta^{1/2}} \left\{ u(y) + \frac{1}{2\delta}|x - y|^2 \right\}.$$

It is then straightforward to see that u^δ and u_δ are respectively decreasing and increasing pointwise as δ decreases to 0, and converge whenever u is continuous at x (and thus a.e.) to $u(x)$.

For fixed $x \in \mathbb{R}^d$, $\delta > 0$, and $\eta > 0$, define

$$A_{\delta,\eta}(x) := \left\{ y \in B_{R\delta^{1/2}}(x) : u(y) - \frac{|x - y|^2}{2\delta} > u^\delta(x) - \eta \right\}.$$

Then, by definition, $A_{\delta,\eta}(x)$ is nonempty, and in fact has nonzero Lebesgue measure. Therefore, for any $x' \in \mathbb{R}^d$ and $y \in A_{\delta,\eta}(x)$, we have

$$u^\delta(x) - u^\delta(x') \leq \frac{|x' - y|^2}{2\delta} - \frac{|x - y|^2}{2\delta} + \eta \leq \frac{R}{\delta^{1/2}}|x' - x| + \frac{|x' - x|^2}{\delta} + \eta.$$

Switching the roles of x and x' and using the fact that η was arbitrary, we see that, for all $x \in \mathbb{R}^d$,

$$\limsup_{x' \to x} \frac{|u^\delta(x') - u^\delta(x)|}{|x' - x|} \leq \frac{R}{\delta^{1/2}}.$$

We may then let R decrease down to $(\operatorname{ess\,sup} u - \operatorname{ess\,inf} u)^{1/2}$, and the same proof for u_δ holds.

For any $\eta > 0$ and a.e. $y \in A_\delta^\eta$,

$$u^{2\delta}(x) \geq u(y) - \frac{|x - y|^2}{4\delta} > u^\delta(x) + \frac{|x - y|^2}{2\delta} - \eta,$$

and so

$$|y - x| \leq 2(u^{2\delta}(x) - u^\delta(x) + \eta)^{1/2}\delta^{1/2}.$$

Therefore, for a.e. y such that $|y - x| > R^\delta(x)$, we must have $u(y) - \frac{|x-y|^2}{2\delta} < u^\delta(x)$, and the statement about restricting the ess sup follows. The corresponding result for u_δ is proved in the same way. $\square$

A formal calculation using the one-sided Lipschitz condition on b suggests that, if u solves (4.2), then the sup- and inf-convolutions of u in the spatial variable are approximate sub- and supersolutions of (4.2). The following result not only establishes this property rigorously, but also proves that it in fact characterizes the unique duality solution of (4.2). The result is proved by using the duality property in relation to a nonnegative distributional solution of (4.1), and we use exactly the same methods to prove the uniqueness of nonnegative distributional solutions in Theorem 4.11 below.

Theorem 4.10. *Assume $u \in C([0, T], L^1_{\mathrm{loc}}(\mathbb{R}^d)) \cap L^\infty([0, T] \times \mathbb{R}^d)$ is continuous almost everywhere and $u(T, \cdot) = u_T \in L^\infty(\mathbb{R}^d)$. Then u is the duality solution of (4.2) if and only if there exist $r^\delta, r_\delta \in L^1_{\mathrm{loc}}([0, T] \times \mathbb{R}^d))$ such that $\lim_{\delta \to 0} r^\delta = \lim_{\delta \to 0} r_\delta = 0$ in L^1_{loc}, and the sup- and inf-convolutions*

$$u^\delta(t, x) := \operatorname*{ess\,sup}_{y \in \mathbb{R}^d} \left\{ u(t, y) - \frac{1}{2\delta}|x - y|^2 \right\}$$

and

$$u_\delta(t, x) := \operatorname*{ess\,inf}_{y \in \mathbb{R}^d} \left\{ u(t, y) + \frac{1}{2\delta}|x - y|^2 \right\}$$

satisfy in the sense of distributions on $[0, T] \times \mathbb{R}^d$ the inequalities

$$\frac{\partial u^\delta}{\partial t} + b(t, x) \cdot \nabla u^\delta \leq r^\delta(t, x) \quad and \quad \frac{\partial u_\delta}{\partial t} + b(t, x) \cdot \nabla u_\delta \geq -r_\delta(t, x).$$

Proof. Assume first that the sup- and inf-convolutions have the stated properties. For standard mollifiers $(\rho_\eta)_{\eta > 0}$ on $\mathbb{R}$, define $u^\delta_\eta(t, x) = (u^\delta(\cdot, x) *_t \rho_\eta)(t)$ and $u_{\delta,\eta}(t, x) = (u_\delta(\cdot, x) *_t \rho_\eta)(t)$. Then, by Lemma 4.3, u^δ_η and $u_{\delta,\eta}$ are Lipschitz continuous on $[0, T] \times \mathbb{R}^d$, and satisfy a.e. in $[0, T] \times \mathbb{R}^d$

$$\frac{\partial u^\delta_\eta}{\partial t} + b(t, x) \cdot \nabla u^\delta_\eta \leq r^\delta_\eta(t, x) \quad \text{and} \quad \frac{\partial u_{\delta,\eta}}{\partial t} + b(t, x) \cdot \nabla u_{\delta,\eta} \geq -r_{\delta,\eta}(t, x),$$

where

$$r^\delta_\eta(t, x) = (r^\delta(\cdot, x) *_t \rho_\eta)(t) + \int_{\mathbb{R}} (b(t, x) - b(s, x)) \cdot \nabla u^\delta(s, x) \rho_\eta(s - t) ds$$

and

$$r_{\delta,\eta}(t, x) = (r_\delta(\cdot, x) *_t \rho_\eta)(t) + \int_{\mathbb{R}} (b(t, x) - b(s, x)) \cdot \nabla u_\delta(s, x) \rho_\eta(s - t) ds.$$

The (local) boundedness of b, ∇u_δ, and ∇u^δ then allows us to invoke the dominated convergence theorem to say that, for fixed δ, $\lim_{\eta \to 0} r^\delta_\eta = r^\delta$ and $\lim_{\eta \to 0} r_{\delta,\eta} = r_\delta$ in L^1_{loc}.

Now let $f_0 \in C_c(\mathbb{R}^d)$ be nonnegative and let f be the BJM solution of (4.1). In view of the nonnegativity of J, f given by (4.3) is nonnegative on $[0, T] \times \mathbb{R}^d$, and the bounds for the backward flow in Lemma 2.2 imply that f has compact support in $[0, T] \times \mathbb{R}^d$. By Theorem 4.1, f is a distributional solution, and therefore

$$\int_{\mathbb{R}^d} f(T, x) u_\eta^\delta(T, x) dx - \int_{\mathbb{R}^d} f_0(x) u_\eta^\delta(0, x) dx$$
$$= \int_0^T \int_{\mathbb{R}^d} f(t, x) \left[\partial_t u^{\delta,\eta}(t, x) + b(t, x) \cdot \nabla u^{\delta,\eta}(t, x) \right] dx dt$$
$$\leq \int_0^T \int_{\mathbb{R}^d} f(t, x) r_\eta^\delta(t, x) dx dt.$$

Sending first $\eta \to 0$ and then $\delta \to 0$, using Lemma 4.3 and the dominated convergence theorem, we conclude that

$$\int_{\mathbb{R}^d} f(T, x) u_T(x) dx \leq \int_{\mathbb{R}^d} f_0(x) u(0, x) dx.$$

Arguing similarly with $u_{\delta,\eta}$ as a test function, we achieve the opposite inequality. By linearity, the duality identity holds for any $f_0 \in L^\infty$ with bounded support, and we conclude that u is the unique duality solution.

Assume now conversely that u is the duality solution. Let $(b^\varepsilon)_{\varepsilon>0}$ be as in (2.3), let u^ε be the corresponding solution, and define

$$u^{\varepsilon,\delta}(t, x) := \sup_{y \in \mathbb{R}^d} \left\{ u^\varepsilon(t, y) - \frac{1}{2\delta} |x - y|^2 \right\}$$

and

$$u_\delta^\varepsilon(t, x) := \inf_{y \in \mathbb{R}^d} \left\{ u^\varepsilon(t, y) + \frac{1}{2\delta} |x - y|^2 \right\}.$$

By Lemma 4.3, for fixed $\delta > 0$, $u^{\varepsilon,\delta}$ and u_δ^ε are Lipschitz continuous in the space variable, uniformly over $[0, T] \times \mathbb{R}^d$ and $\varepsilon > 0$. Moreover, the sup and inf are actually a max and min, and may be restricted to

$$|y - x| \leq (\max u_0 - \min u_0)^{1/2} \delta^{1/2}$$

(note that we have used the maximum principle for the transport equation to control the maximum and minimum of u^ε and u_ε). We may alternatively restrict the y for which the maximum in the definition of $u^{\varepsilon,\delta}(t, x)$ is attained to satisfy

$$|y - x| \leq 2(u^{\varepsilon,2\delta}(t, x) - u^{\varepsilon,\delta}(t, x))^{1/2} \delta^{1/2}, \tag{4.11}$$

and the minimum in the definition of u_δ^ε is attained by y satisfying

$$|y - x| \leq 2(u_\delta^\varepsilon(t, x) - u_{2\delta}^\varepsilon(t, x))^{1/2} \delta^{1/2}. \tag{4.12}$$

Standard properties of envelopes then give the identities, for any $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$\frac{\partial u^{\varepsilon,\delta}}{\partial t}(t, x) = \frac{\partial u^\varepsilon}{\partial t}(t, y) \quad \text{and} \quad \nabla u^{\varepsilon,\delta}(t, x) = \nabla u^\varepsilon(t, y) = \frac{y - x}{\delta}$$

for some y satisfying (4.11). Therefore

$$\partial_t u^{\varepsilon,\delta}(t, x) = -b^\varepsilon(t, y) \cdot \nabla u^{\varepsilon,\delta}(t, x),$$

from which we deduce that $u^{\varepsilon,\delta}$ is uniformly Lipschitz continuous in the time variable over $[0, T] \times B_R$ for any $R > 0$, independently of ε. Further developing the equality gives

$$\begin{aligned}
\frac{\partial u^{\varepsilon,\delta}}{\partial t}(t, x) + b^\varepsilon(t, x) \cdot \nabla u^{\varepsilon,\delta}(t, x) &= \frac{\partial u^\varepsilon}{\partial t}(t, y) + b^\varepsilon(t, x) \cdot \nabla u^\varepsilon(t, y) \\
&= -(b^\varepsilon(t, x) - b^\varepsilon(t, y)) \cdot \frac{x - y}{\delta} \\
&\leq C_0(t) \frac{|x - y|^2}{\delta} \\
&\leq 4C_0(t)(u^{\varepsilon,2\delta}(t, x) - u^{\varepsilon,\delta}(t, x)). \quad (4.13)
\end{aligned}$$

We similarly have that u_δ^ε is Lipschitz continuous in the time variable, locally in space, uniformly over $\varepsilon > 0$, and

$$\frac{\partial u_\delta^\varepsilon}{\partial t}(t, x) + b^\varepsilon(t, x) \cdot \nabla u_\delta^\varepsilon(t, x) \geq -4C_0(t)(u_\delta^\varepsilon(t, x) - u_{2\delta}^\varepsilon(t, x)). \quad (4.14)$$

We now claim that, as $\varepsilon \to 0$, $u^{\varepsilon,\delta}$ and u_δ^ε converge pointwise to respectively u^δ and u_δ, and then, by the uniform-in-ε Lipschitz regularity, the convergence is locally uniform. To see this, fix $x \in \mathbb{R}^d$ and $\eta > 0$, and let $A \subset \mathbb{R}^d$ be a set of positive measure such that

$$u^\delta(t, x) \leq u(t, y) - \frac{|x - y|^2}{2\delta} + \eta.$$

We then have, for all $y \in A$,

$$u^{\delta,\varepsilon}(t, x) \geq u^\varepsilon(t, y) - \frac{|x - y|^2}{2\delta}.$$

For at least one such y, we then have $u^\varepsilon(t, y) \xrightarrow{\varepsilon \to 0} u(t, y)$, and we thus have

$$\limsup_{\varepsilon \to 0} \left(u^\delta(t, x) - u^{\delta,\varepsilon}(t, x) \right) \leq \eta.$$

It follows that $\limsup_{\varepsilon \to 0} \left(u^\delta(t, x) - u^{\delta,\varepsilon}(t, x) \right) \leq 0$ since η was arbitrary.

Now, there exists a full measure set $B \subset \mathbb{R}^d$ such that, for all $y \in B$,

$$u^\delta(t, x) \geq u(t, y) - \frac{|x - y|^2}{2\delta} \quad \text{and} \quad \lim_{\varepsilon \to 0} u^\varepsilon(t, y) = u(t, y).$$

In view of the continuity of $u^\varepsilon(t, \cdot)$, there exists a bounded (independently ε) sequence $(y_n)_{n \in \mathbb{N}} \subset B$ such that

$$\rho_n := u^{\delta, \varepsilon}(t, x) - \left\{ u^\varepsilon(t, y_n) - \frac{|x - y_n|^2}{2\delta} \right\}$$

satisfies $\lim_{n \to \infty} \rho_n = 0$. Therefore, for all n,

$$u^{\delta, \varepsilon}(t, x) - u^\delta(t, x) \le u^\varepsilon(t, y_n) - u(t, y_n) + \rho_n.$$

Sending $\varepsilon \to 0$ gives $\limsup_{\varepsilon \to 0} (u^{\delta, \varepsilon}(t, x) - u^\delta(t, x)) \le \rho_n$, and the proof of pointwise convergence is finished upon sending $n \to \infty$. The exact same argument can be used for the pointwise convergence of u^ε_δ to u_δ.

It then follows that, for fixed δ, as $\varepsilon \to 0$, $\nabla u^{\varepsilon, \delta}$ and $\nabla u^\varepsilon_\delta$ converge weak-$\star$ in L^∞ to ∇u^δ and ∇u_δ respectively, while b^ε converges in L^1_{loc} to b. We may then take $\varepsilon \to 0$ in (4.13) and (4.14) to obtain the distributional inequalities

$$\frac{\partial u^\delta}{\partial t}(t, x) + b(t, x) \cdot \nabla u^\delta(t, x) \le 4C_0(t)(u^{2\delta}(t, x) - u^\delta(t, x)) =: r^\delta(t, x)$$

and

$$\frac{\partial u_\delta}{\partial t}(t, x) + b(t, x) \cdot \nabla u_\delta(t, x) \ge -4C_0(t)(u_\delta(t, x) - u_{2\delta}(t, x)) =: -r_\delta(t, x).$$

By Lemma 4.3 and the almost-everywhere continuity of u, the right-hand sides of both inequalities converge a.e. to 0 as $\delta \to 0$, and, by the uniform boundedness in δ of u^δ and u_δ and the dominated convergence theorem, r^δ and r_δ both converge in L^1_{loc} to 0 as $\delta \to 0$. $\qquad\qquad\square$

4.4.2. The Conservative Equation: Uniqueness of Nonnegative Solutions
We observe that, in the first implication in the proof of Theorem 4.10, it was proved that u was a duality solution by proving the duality identity relative to a "good" nonnegative solution, i.e. the reversible BJM solution we have been working with above. However, it was only explicitly used that f was a distributional solution. Therefore, after having proved the equivalence in Theorem 4.10, we arrive at the following:

Theorem 4.11. *Suppose that $f \in C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$ is a distributional solution of (4.1) and $f \ge 0$. Then $f(t, x) = f(0, \phi_{0,t}(x)) J_{0,t}(x)$.*

Proof. Fix $t > 0$ and $v \in C_c(\mathbb{R}^d)$, and let $u \in C([0, t], L^1_{\text{loc}}(\mathbb{R}^d)) \cap L^\infty([0, t] \times \mathbb{R}^d)$ be the duality solution of (4.2) with terminal data v at time t. Then, by Theorem 4.7, u is continuous almost everywhere in $[0, t] \times \mathbb{R}^d$. Arguing exactly as in the first part of Theorem 4.10, using the nonnegativity of f, we arrive at the equality

$$\int_{\mathbb{R}^d} f(t, x) v(x) dx = \int_{\mathbb{R}^d} f(0, x) u(0, x) dx.$$

Since v was arbitrary, it follows from the definition of duality solutions that $f(t, x)$ must be given by (4.3). $\qquad\qquad\square$

We then have the following corollary about characterizing the BJM solution even when f is signed:

Corollary 4.3. *A function* $f \in C([0, T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$ *is the unique reversible solution of* (4.1) *in the sense of* [25] *if and only if* f *and* $|f|$ *are both solutions in the sense of distributions.*

Proof. That this property is satisfied by the good solution was already pointed out (Corollary 4.1). Suppose now that f and $|f|$ are both distributional solutions. It follows that $f_+ = \frac{1}{2}(f + |f|)$ and $f_- = \frac{1}{2}(|f| - f)$ are distributional solutions, and, since $f_+ \geq 0$ and $f_- \geq 0$, they are both BJM reversible solutions. Therefore $f = f_+ - f_-$ is a reversible solution by linearity.   □

4.4.3. Uniqueness of Regular Lagrangian Flows   We can finally establish the uniqueness for the forward flows of the ODE (2.2)

Theorem 4.12. *For every* $s \in [0, T]$ *and almost every* $x \in \mathbb{R}^d$, $\phi_{t,s}(x)$ *is the unique absolutely continuous solution of* (2.2).

Proof. This is a consequence of Theorem 4.11 and the superposition principle of Ambrosio [6, Theorem 3.1].   □

4.5. Some Remarks for Second Order Equations

We next investigate the second-order analogues of (4.1) and (4.2). As mentioned earlier, we are not able to treat the most general case in which σ is a regular function of x. This is due to the fact that Lemma 2.5 only gives regularity of the backward stochastic flow in $C^{0,1-\varepsilon}$ for $0 < \varepsilon < 1$. As a consequence, defining the Jacobian and using it to analyze the right-inverse of the flow is not possible at present with our methods. Our results in this case are limited to stochastic flows for which the coefficient σ in front of the Wiener process is constant in the space variable. The generalization to regular but nonconstant σ will be the subject of future work.

4.5.1. The Expansive Stochastic Flow with Constant Noise Coefficient   The stochastic analogue of the forward flow (2.2) is

$$d_t \Phi_{t,s}(x) = b(t, \Phi_{t,s}(x))dt + \sigma(t, \Phi_{t,s}(x))dW_t, \quad t \in [s, T], \quad \Phi_{s,s}(x) = x, \tag{4.15}$$

where $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$ is some matrix-valued map. As we shall see, this general setting is out of the reach at the moment, and we thus assume

$$\sigma \in L^2([0, T], \mathbb{R}^{d \times m}) \tag{4.16}$$

is constant in the space variable. We then consider the forward stochastic flow

$$d\Phi_{t,s}(x) = b(t, \Phi_{t,s}(x))dt + \sigma_t dW_t, \quad t \in [s, T], \quad \Phi_{s,s}(x) = x. \tag{4.17}$$

Formally defining

$$\tilde{\Phi}_{t,s}(x) := \Phi_{t,s}(x) - \underbrace{\int_s^t \sigma_r dW_r}_{:=M_t - M_s}$$

leads to the random ODE

$$\partial_t \tilde{\Phi}_{t,s}(x) = b\left(t, \tilde{\Phi}_{t,s}(x) + M_t - M_s\right), \quad t \in [s, T], \quad \tilde{\Phi}_{s,s}(x) = x. \quad (4.18)$$

We now invoke the theory of the previous subsections to obtain the following:

Theorem 4.13. *For every $s \in [0, T)$, with probability one, there exists a unique $\Phi_{\cdot,s} \in C([s, T], L^p_{\mathrm{loc}}(\mathbb{R}^d)) \cap L^p_{\mathrm{loc}}(\mathbb{R}^d, C([s, T]))$ such that, for a.e. $x \in \mathbb{R}^d$,*

$$\Phi_{t,s}(x) = x + \int_s^t b(r, \Phi_{r,s}(x))dr + \int_s^t \sigma_r dW_r.$$

If $(b^\varepsilon)_{\varepsilon > 0}$ are as in (2.3) and Φ^ε is the unique stochastic flow solving (4.17) with drift b^ε, then, with probability one, as $\varepsilon \to 0$, Φ^ε converges in $C([s, T], L^p_{\mathrm{loc}}(\mathbb{R}^d))$ and in $L^p_{\mathrm{loc}}(\mathbb{R}^d, C([s, T]))$ to Φ.

Proof. This follows upon applying the results of Theorems 4.8 and 4.12 to the random ODE (4.18). □

4.5.2. A Priori Estimates for the Second-Order Nonconservative Equation

We next relate the forward stochastic flow from the previous subsection to the terminal value problem for a certain second-order, nonconservative equation. This will be done with the use of a priori L^p and BV estimates, which lead to useful compactness results, just as for the first order case.

We begin with the more general problem

$$-\partial_t u - \mathrm{tr}[a(t, x)\nabla^2 u] + b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T, \quad (4.19)$$

where

$$a(t, x) = \frac{1}{2}\sigma(t, x)\sigma(t, x)^T, \quad \sigma \in L^2([0, T], C^{1,1}(\mathbb{R}^d, \mathbb{R}^{d \times m})); \quad (4.20)$$

notice that, although we allow σ to be nonconstant here, we require more regularity for σ than in Sect. 3.

Lemma 4.4. *There exists $C \in L^1_+([0, T])$ depending only on the $C^{1,1}$ norm of σ such that, if u is a smooth solution of*

$$-\partial_t u - \mathrm{tr}[a(t, x)\nabla^2 u] = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T,$$

then

$$\|u(t, \cdot)\|_{BV(\mathbb{R}^d)} \leq \exp\left(\int_t^T C(s)ds\right) \|u_T\|_{BV(\mathbb{R}^d)}.$$

Proof. For $(t, x, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$, set $w(t, x, z) = \nabla u(t, x) \cdot z$. Then w solves the parabolic PDE

$$\frac{\partial w}{\partial t} - \text{tr}[A(t, x, z)\nabla^2_{(x,z)}w] = 0 \quad \text{in } (0, T) \times \mathbb{R}^{2d},$$

where

$$A(t, x, z) = \frac{1}{2}\begin{pmatrix} \sigma(t, x) \\ z \cdot \nabla\sigma(t, x) \end{pmatrix}\left(\sigma(t, x)^T \ z \cdot \nabla\sigma(t, x)^T\right).$$

After a routine regularization argument, using the convexity of $w \mapsto |w|$, we find that

$$\frac{\partial |w|}{\partial t} - \text{tr}[A(t, x, z)\nabla^2_{(x,z)}|w|] \leq 0 \quad \text{in } (0, T) \times \mathbb{R}^d \times \mathbb{R}^d. \tag{4.21}$$

For some $m > d + 1$, let $\phi \in C^\infty_+([0, \infty))$ be such that, for some universal $C > 0$,

$$\phi(r) = \frac{1}{r^m} \quad \text{for } r \geq 1 \quad \text{and} \quad r|\phi'(r)| + r^2|\phi''(r)| \leq C\phi(r) \quad \text{for all } r \geq 0. \tag{4.22}$$

We multiply (4.21) by $\phi(|z|)$ and integrate in $(x, z) \in \mathbb{R}^d \times \mathbb{R}^d$. Then (4.20) and (4.22) imply that for some $C \in L^1_+([0, T])$,

$$-\frac{d}{dt}\iint_{\mathbb{R}^d \times \mathbb{R}^d} |w(t, x, z)|\phi(|z|)dxdz \leq C(t)\iint_{\mathbb{R}^d \times \mathbb{R}^d} |w(t, x, z)|\phi(|z|)dxdz.$$

The proof is then finished by Grönwall's lemma and the fact that

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} |w(t, x, z)|\phi(z)dxdz = c_0 \int_{\mathbb{R}^d} |\nabla u(t, x)|dx,$$

where $c_0 := \int_{\mathbb{R}^d} |v \cdot z|\phi(|z|)dz$ is finite and independent of $|v| = 1$. $\square$

We have already proved an exponential propagation of the BV bounds when $a = 0$ in Lemma 4.1. It is a classical fact for evolution PDEs that, upon using a splitting scheme, that these estimates can be combined, and we immediately have the following:

Lemma 4.5. *There exists a constant $C \in L^1_+([0, T])$ depending only on the constants in* (2.1) *and* (4.20) *such that, if u is a smooth solution of* (4.19)*, then*

$$\|u(t, \cdot)\|_{L^p} \leq \exp\left(\int_0^t C(s)ds\right)\|u_T\|_{L^p}$$

and

$$\|u(t, \cdot)\|_{BV} \leq \exp\left(\int_0^t C(s)ds\right)\|u_T\|_{BV}.$$

Just as in the first-order case, it is not possible to define L^p-distributional solutions of (4.19), and the utility of Lemma 4.5 is that it allows to obtain strongly convergent subsequences in $C([0, T], L^p(\mathbb{R}^d))$ after regularizing the velocity field b.

The main question is whether such limiting solutions are unique. This uniqueness was achieved in the first-order case through duality with the conservative equation, and the solution was further characterized with a formula involving the forward flow. In the second-order case, we are constrained to work with constant noise coefficients:

$$-\partial_t u - \mathrm{tr}[a(t)\nabla^2 u] + b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T, \quad (4.23)$$

where $a = \frac{1}{2}\sigma\sigma^T$ as before.

Theorem 4.14. *For $1 < p < \infty$ and $t \in [0, T]$, the map*

$$C_c(\mathbb{R}^d) \ni u_T \mapsto \mathbb{E}[u_T \circ \Phi_{T,t}]$$

extends to a continuous, linear, order-preserving map on $L^p(\mathbb{R}^d)$, and the function

$$u(t, x) := \mathbb{E}[u_T(\Phi_{T,t}(x))] \quad (t, x) \in [0, T] \times \mathbb{R}^d \qquad (4.24)$$

belongs to $C([0, T], L^p(\mathbb{R}^d))$, and, if $u_T \in BV(\mathbb{R}^d)$, then $u \in L^\infty([0, T], BV(\mathbb{R}^d))$.

If $(b^\varepsilon)_{\varepsilon > 0}$ is as in (2.3) and u^ε is the corresponding solution of (4.23), then, as $\varepsilon \to 0$, u^ε converges strongly to u in $C([0, T], L^p(\mathbb{R}^d))$.

Proof. Assume that $u_T \in C^2(\mathbb{R}^d) \cap C_c(\mathbb{R}^d)$. For b^ε and u^ε as in the statement of the theorem, we have the standard representation formula $u^\varepsilon(t, x) = \mathbb{E}[u_T(\Phi^\varepsilon_{T,t}(x))]$, where Φ^ε corresponds to the flow (4.17) with drift b^ε. By Theorem 4.13, for any $t \in [0, T]$, with probability one, $u_T \circ \Phi^\varepsilon_{T,t} \to u_T \circ \Phi_{T,t}$ a.e. in $\mathbb{R}^d$. On the other hand, by Lemma 4.5, $(u^\varepsilon)_{\varepsilon > 0}$ is precompact in $C([0, T], L^p(\mathbb{R}^d))$, and therefore the full sequence converges to u given by (4.24). The L^p-bounds and the extension to $u_T \in L^p(\mathbb{R}^d)$ now follow from the L^p a priori estimates in Lemma 4.5. $\quad\square$

4.5.3. Representation Formula for the Fokker–Planck Equation

We turn next to the Fokker-Planck equation

$$\partial_t f - \nabla^2 \cdot (a(t, x) f) + \mathrm{div}(b(t, x) f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad f(0, \cdot) = f_0, \quad (4.25)$$

where once again $a = \frac{1}{2}\sigma\sigma^T$ with σ as in (4.20).

The existence of solutions in $C([0, T], L^p(\mathbb{R}^d))$ is straightforward; we include the proof for convenience.

Theorem 4.15. *For any $f_0 \in L^p(\mathbb{R}^d)$, $1 \leq p \leq \infty$, there exists a distributional solution $f \in C([0, T], L^p_{\mathrm{w}}(\mathbb{R}^d))$ if $1 \leq p < \infty$, or $f \in L^\infty$ if $p = \infty$. Moreover, there exists $C \in L^1_+([0, T])$ depending only on p, $C_0(t)$ from (2.1) and the $L^2([0, T], C^{1,1}(\mathbb{R}^d))$ norm of a[8] such that*

$$\|f(t, \cdot)\|_{L^p} \leq \exp\left(\int_0^t C(s)ds\right) \|f\|_{L^p}.$$

---

[8] In fact, only an upper bound for $\nabla^2 \cdot a = \partial_{ij} a_{ij}$ is needed.

Proof. We do this with the use of a priori estimates, assuming all the data is smooth. The computations may be made rigorous by regularizing b, adding a small ellipticity to a, and extracting weakly convergent subsequences.

We then compute

$$\partial_t |f|^p - \nabla^2 \cdot (a(t,x)|f|^p) + \mathrm{div}(b(t,x)|f|^p)$$
$$\leq (p-1)\left(\nabla^2 \cdot a(t,x) - \mathrm{div}\, b(t,x)\right)|f|^p,$$

and so $\partial_t \int |f(t,\cdot)|^p \leq C(t) \int |f(t,\cdot)^p$ for some C as in the statement of the Theorem. The result now follows from Grönwall's lemma. $\qquad\square$

We now explore the possibility of obtaining a formula for the solution, similar to (4.3) for the first order equation (4.1). To do so, it is convenient to reverse time and consider, for fixed $t \in (0,T]$, the equation satisfied by $g^{(t)}(s,x) := f(t-s,x)$:

$$-\partial_s g^{(t)} - \nabla^2 \cdot (a(t-s,x)g^{(t)}) + \mathrm{div}(b(t-s,x)g^{(t)}) = 0 \quad \text{in } (0,t) \times \mathbb{R}^d,$$
$$g^{(t)}(t,\cdot) = f_0.$$

For $(s,x,\xi) \in [0,t] \times \mathbb{R}^d \times \mathbb{R}$, define $G^{(t)}(s,x,\xi) = g^{(t)}(s,x)\xi$. Then

$$
\begin{cases}
-\partial_s G^{(t)} - \mathrm{tr}[A^{(t)}(s,x,\xi)\nabla^2_{x,\xi} G^{(t)}] \\
\quad - B^{(t)}(s,x) \cdot \nabla G^{(t)} - C^{(t)}(s,x)\xi \partial_\xi G^{(t)} = 0 \quad \text{in } (0,t) \times \mathbb{R}^{d+1}, \quad (4.26) \\
G^{(t)}(t,x,\xi) = f_0(x)\xi,
\end{cases}
$$

where

$$
\begin{cases}
A^{(t)}(s,x,\xi) = \dfrac{1}{2}\Sigma^{(t)}(s,x,\xi)\Sigma^{(t)}(s,x,\xi)^T, \quad \Sigma^{(t)}(s,x,\xi) = \begin{pmatrix} \sigma \\ \xi\, \mathrm{div}\, \sigma \end{pmatrix}, \\[2mm]
B^{(t)}(s,x) = -b + (\sigma \cdot \nabla)\sigma^T, \quad \text{and} \\[2mm]
C^{(t)}(s,x) = -\mathrm{div}\,(b - \mathrm{div}\, a) \\[2mm]
\qquad\qquad = -\mathrm{div}\, b + \mathrm{tr}[(\sigma \cdot \nabla)(\nabla \cdot \sigma)] + \dfrac{1}{2}|\mathrm{div}\,\sigma|^2 + \dfrac{1}{2}\mathrm{tr}[\nabla\sigma\nabla\sigma^T];
\end{cases}
$$
$$(4.27)$$

for brevity, we have suppressed the arguments for a, σ, and b, which are all $(t-s,x)$.

For an m-dimensional Wiener process W on $[0,t]$ and a fixed $s \in [0,t]$, we are led to consider the SDE, for $r \in [s,t]$,

$$
\begin{cases}
d_r \begin{pmatrix} \Phi_{r,s}^{(t)}(x,\xi) \\ \Xi_{r,s}^{(t)}(x,\xi) \end{pmatrix} = \begin{pmatrix} B^{(t)}(r, \Phi_{r,s}^{(t)}(x,\xi)) \\ C^{(t)}(r, \Phi_{r,s}^{(t)}(x,\xi))\Xi_{r,s}^{(t)}(x,\xi) \end{pmatrix} dr \\[3mm]
\qquad + \Sigma^{(t)}(r, \Phi_{r,s}^{(t)}(x,\xi), \Xi_{r,s}^{(t)}(x,\xi))dW_r, \quad (4.28) \\[3mm]
\begin{pmatrix} \Phi_{s,s}^{(t)}(x,\xi) \\ \Xi_{s,s}^{(t)}(x,\xi) \end{pmatrix} = \begin{pmatrix} x \\ \xi \end{pmatrix}.
\end{cases}
$$

Itô's formula, (4.26), and (4.28) then yield that, for any $(s,x,\xi) \in [0,t) \times \mathbb{R}^d \times \mathbb{R}$,

$$r \mapsto G^{(t)}(r, \Phi_{r,s}^{(t)}(x,\xi), \Xi_{r,s}^{(t)}(x,\xi))$$

is a martingale on $[s, t]$ with respect to the filtration $(\mathcal{F}_r)_{r \in [0,t]}$ generated by the Wiener process W, and so, for all $r \in [s, t]$,

$$\mathbb{E}\left[G^{(t)}(r, \Phi_{r,s}^{(t)}(x, \xi), \Xi_{r,s}^{(t)}(x, \xi)) \mid \mathcal{F}_s\right] = G^{(t)}(s, x, \xi). \qquad (4.29)$$

Observe that $\Phi_{r,s}^{(t)}$ is independent of ξ, while $\Xi_{r,s}^{(t)}$ can be written as $\Xi_{r,s}^{(t)}(x, \xi) = J_{r,s}^{(t)}(x)\xi$ for some scalar quantity $J_{r,s}^{(t)}(x)$, and so (4.28) reduces to the two SDEs

$$\begin{cases} d_r \Phi_{r,s}^{(t)}(x) = -\left[b(t - r, \Phi_{r,s}^{(t)}(x)) - (\sigma \cdot \nabla)\sigma^T(t - r, \Phi_{r,s}^{(t)}(x))\right] dt \\ \qquad\qquad + \sigma(t - r, \Phi_{r,s}^{(t)}(x)) dW_r, \quad r \in [s, t], \\ \Phi_{s,s}^{(t)}(x) = x \end{cases} \qquad (4.30)$$

and

$$\begin{cases} d_r J_{r,s}^{(t)}(x) = \left[-\operatorname{div} b + \operatorname{tr}[(\sigma \cdot \nabla)(\nabla \cdot \sigma)] + \frac{1}{2}|\operatorname{div} \sigma|^2 \right. \\ \qquad\qquad \left. + \frac{1}{2}\operatorname{tr}[\nabla\sigma\nabla\sigma^T] \right](t - r, \Phi_{r,s}^{(t)}(x)) J_{r,s}^{(t)}(x) dr \\ \qquad\qquad + \operatorname{div}\sigma(t - r, \Phi_{r,s}^{(t)}(x)) J_{r,s}^{(t)}(x) dW_r, \quad r \in [s, t], \\ J_{s,s}^{(t)}(x) = 1. \end{cases} \qquad (4.31)$$

Standard but tedious computations involving Itô's formula reveal that $J_{r,s}^{(t)}(x) = \det \nabla_x \Phi_{r,s}^{(t)}(x)$.

Taking $r = t$ and $\xi = 1$ in (4.29), we thus arrive at

$$\mathbb{E}\left[f_0(\Phi_{t,s}^{(t)}(x)) J_{t,s}^{(t)}(x) \mid \mathcal{F}_s\right] = g(s, x),$$

and so, because $g(0, x) = f(t, x)$, we obtain the representation for solutions of (4.25):

$$f(t, x) = \mathbb{E}\left[f_0(\Phi_{t,0}^{(t)}(x)) J_{t,0}^{(t)}(x)\right]. \qquad (4.32)$$

Let us note that $\Phi_{t,0}^{(t)}$ has the same law as $(\Phi_{t,0})^{-1}$, where $\Phi_{t,s}$ is the stochastic flow from (4.15). We can see this by duality with the nonconservative equation. Indeed, if u is the solution of (4.19) with $u(t, \cdot) = g$ for some given g, then

$$\int f_0(x) u(0, x) dx = \int f(t, x) g(x) dx.$$

On the other hand, by (4.24) and (4.32),

$$\int f_0(x) u(0, x) dx = \mathbb{E}\int f_0(x) g(\Phi_{t,0}(x)) dx$$

and

$$\int f(t, x) g(x) dx = \mathbb{E}\int f_0(\Phi_{t,0}^{(t)}(x)) g(x) J_{t,0}^{(t)}(x) dx,$$

so, using the change of variables formula and the fact that f_0 is arbitrary, we have $\mathbb{E}[g(\Phi_{t,0}(x))] = \mathbb{E}[g([\Phi_{t,0}^{(t)}]^{-1}(x))]$ for all $g : \mathbb{R}^d \to \mathbb{R}$ and $x \in \mathbb{R}^d$.

We now note that the SDE (4.30) falls under the assumptions of Lemma 2.5, and therefore, for every $0 \leq s < t \leq T$, there exists a unique solution $\Phi_{\cdot,s}^{(t)}$ with the properties laid out by that result. However, the main difficulty is that we do not know whether $\Phi_{t,0}^{(t)}$ is Lipschitz continuous on $\mathbb{R}^d$ (see Remark 2.8). This prevents us from bounding $J_{t,0}^{(t)}$ uniformly in L^∞ and passing to weak distributional limits. This is a major obstacle in using the formula (4.32) to identify the unique limiting distributional solution of (4.25), as we did for the first order equation (4.1).

The exception is when σ is independent of x. In that case, (4.30) and (4.31) become

$$d_r \Phi_{r,s}^{(t)}(x) = -b(t - r, \Phi_{r,s}^{(t)}(x))dr + \sigma(t - r)dW_r, \quad r \in [s, t], \quad \Phi_{s,s}^{(t)}(x) = x \tag{4.33}$$

and

$$\partial_r J_{r,s}^{(t)}(x) = -\operatorname{div} b(t - r, \Phi_{r,s}^{(t)}(x))J_{r,s}^{(t)}(x), \quad r \in [s, t], \quad J_{s,s}^{(t)}(x) = 1. \tag{4.34}$$

The SDE (4.34) is in fact an ODE with random coefficients. In particular, $J_{\cdot,s}^{(t)}$ has a deterministic bound.

We then characterize uniquely the limiting distributional solution of

$$\partial_t f - \nabla^2 \cdot (a(t)f) + \operatorname{div}(b(t, x)f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad f(0, \cdot) = f_0. \tag{4.35}$$

Theorem 4.16. *For $1 \leq p < \infty$, the formula (4.32), where $\Phi_{\cdot,s}^{(t)}$ and $J_{\cdot,s}^{(t)}$ are specified by respectively (4.33) and (4.34), extends continuously to any $f_0 \in L^p(\mathbb{R}^d)$. If $f_0 \in L^p(\mathbb{R}^d)$ and $(b^\varepsilon)_{\varepsilon > 0}$ are as in (2.3) and f^ε is the corresponding solution of (4.35), then, as $\varepsilon \to 0$, f^ε converges weakly in $C([0, T], L_w^p(\mathbb{R}^d))$ to f. If $f_0 \geq 0$, then there exists a unique nonnegative distributional solution of (4.35), which is given by (4.32).*

Proof. Let $(b^\varepsilon)_{\varepsilon > 0}$ and f^ε be as in the statement of the theorem, and assume $f_0 \in C_c^2(\mathbb{R}^d)$. Let u^ε be the solution of (4.23) with velocity b^ε and with terminal data $u(t, \cdot) = g \in C_c^2(\mathbb{R}^d)$ for some fixed $t \in [0, T]$. Then integration by parts yields

$$\int f^\varepsilon(t, x)g(x)dx = \int f_0(x)u^\varepsilon(0, x)dx.$$

By Theorem 4.14, as $\varepsilon \to 0$, u^ε converges strongly in $L^{p'}(\mathbb{R}^d)$ to the function u defined uniquely by $u(s, x) = g(\Phi_{t,s}(x))$. Therefore, any $C([0, T], L_w^p(\mathbb{R}^d))$-weak limit f of f^ε as $\varepsilon \to 0$ must satisfy

$$\int f(t, x)g(x)dx = \int f_0(x)u(0, x)dx,$$

and it follows that there is a unique such limiting function f.

On the other hand, for $\varepsilon > 0$,

$$f^\varepsilon(t, x) = \mathbb{E}\left[f_0\left(\Phi_{t,0}^{(t),\varepsilon}(x) \right) J_{t,0}^{(t),\varepsilon}(x) \right],$$

where $\Phi_{\cdot,s}^{(t),\varepsilon}$ and $J_{\cdot,s}^{(t),\varepsilon}$ are as in respectively (4.33) and (4.34) with b replaced everywhere by b^ε. For fixed $t \in [0, T]$, uniformly in ε, $\Phi_{t,0}^{(t),\varepsilon}$ is Lipschitz continuous on $\mathbb{R}^d$, and so $J_{t,0}^{(t),\varepsilon} = \det \nabla_x \Phi_{t,0}^{(t),\varepsilon}$ is bounded in L^∞. By exactly the same arguments as in Lemma 2.3 and Theorem 4.1, we see that, as $\varepsilon \to 0$, $\mathbb{E} f_0\left(\Phi_{t,0}^{(t),\varepsilon} \right) J_{t,0}^{(t),\varepsilon}$ converges weakly in L^p to $\mathbb{E} f_0\left(\Phi_{t,0}^{(t)} \right) J_{t,0}^{(t)}$. It follows that f must be given by (4.32). The fact that the formula extends to arbitrary $f_0 \in L^p(\mathbb{R}^d)$ now follows from the a priori L^p bounds in Theorem 4.15.

The uniqueness of nonnegative distributional solutions is then a consequence of the uniqueness of the forward flow established in Theorem 4.13, as well as the generalization of superposition to second-order Fokker–Planck equations (see Figalli [41, Lemma 2.3]). □

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Code availability Not applicable.

Declarations

Conflict of interest P.-L. Lions is a member of the Editorial Board. B. Seeger has no relevant competing interests to declare.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. ALBERTI, G., AMBROSIO, L.: A geometrical approach to monotone functions in $\mathbf{R}^n$. *Math. Z.* **230**(2), 259–316, 1999. https://doi.org/10.1007/PL00004691

2. ALBERTI, G., BIANCHINI, S., CRIPPA, G.: Two-dimensional transport equation with Hamiltonian vector fields. In: Hyperbolic problems: theory, numerics and applications, Proc. Sympos. Appl. Math., vol 67, Part 2. Amer. Math. Soc., Providence, RI, p 337–346, 2009. https://doi.org/10.1090/psapm/067.2/2605229

3. ALBERTI, G., BIANCHINI, S., CRIPPA, G.: Structure of level sets and Sard-type properties of Lipschitz maps. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **12**(4), 863–902, 2013

4. ALBERTI, G., BIANCHINI, S., CRIPPA, G.: A uniqueness result for the continuity equation in two dimensions. *J. Eur. Math. Soc.* **16**(2), 201–234, 2014. https://doi.org/10.4171/JEMS/431

5. AMBROSIO, L.: Transport equation and Cauchy problem for BV vector fields. *Invent. Math.* **158**(2), 227–260, 2004. https://doi.org/10.1007/s00222-004-0367-2

6. AMBROSIO, L.: Transport equation and Cauchy problem for non-smooth vector fields. In: Calculus of variations and nonlinear partial differential equations, Lecture Notes in Math., vol 1927. Springer, Berlin, p 1–41, 2008. https://doi.org/10.1007/978-3-540-75914-0_1, https://doi-org.ezproxy.lib.utexas.edu/10.1007/978-3-540-75914-0_1

7. AMBROSIO, L.: Well posedness of ODE's and continuity equations with nonsmooth vector fields, and applications. *Rev. Mat. Complut.* **30**(3), 427–450, 2017. https://doi.org/10.1007/s13163-017-0244-3

8. AMBROSIO, L., CRIPPA, G.: Continuity equations and ODE flows with non-smooth velocity. *Proc. R. Soc. Edinburgh Sect. A* **144**(6), 1191–1244, 2014. https://doi.org/10.1017/S0308210513000085

9. AMBROSIO, L., DE LELLIS, C.: Existence of solutions for a class of hyperbolic systems of conservation laws in several space dimensions. *Int. Math. Res. Not.* **41**, 2205–2220, 2003. https://doi.org/10.1155/S1073792803131327

10. AMBROSIO, L., TREVISAN, D.: Lecture notes on the DiPerna-Lions theory in abstract measure spaces. *Ann. Fac. Sci. Toulouse Math. (6)* **26**(4), 729–766, 2017. https://doi.org/10.5802/afst.1551

11. AMBROSIO, L., BOUCHUT, F., DE LELLIS, C.: Well-posedness for a class of hyperbolic systems of conservation laws in several space dimensions. *Comm. Partial Differ. Equ.* **29**(9–10), 1635–1651, 2004. https://doi.org/10.1081/PDE-200040210

12. AMBROSIO, L., CRIPPA, G., MANIGLIA, S.: Traces and fine properties of a BD class of vector fields and applications. *Ann. Fac. Sci. Toulouse Math. (6)* **14**(4), 527–561, 2005

13. AMBROSIO, L., LECUMBERRY, M., MANIGLIA, S.: Lipschitz regularity and approximate differentiability of the DiPerna-Lions flow. *Rend. Sem. Mat. Univ. Padova* **114**, 29–50, 2005

14. AMBROSIO, L., DE LELLIS, C., MALÝ, J.: On the chain rule for the divergence of BV-like vector fields: applications, partial results, open problems. In: Perspectives in nonlinear partial differential equations, Contemp. Math., vol 446. Amer. Math. Soc., Providence, RI, p 31–67, 2007. https://doi.org/10.1090/conm/446/08625

15. AMBROSIO, L., COLOMBO, M., DE PHILIPPIS, G., et al.: Existence of Eulerian solutions to the semigeostrophic equations in physical space: the 2-dimensional periodic case. *Comm. Partial Differ. Equ.* **37**(12), 2209–2227, 2012. https://doi.org/10.1080/03605302.2012.669443

16. AMBROSIO, L., COLOMBO, M., DE PHILIPPIS, G., et al.: A global existence result for the semigeostrophic equations in three dimensional convex domains. *Discrete Contin. Dyn. Syst.* **34**(4), 1251–1268, 2014. https://doi.org/10.3934/dcds.2014.34.1251

17. AMBROSIO, L., COLOMBO, M., FIGALLI, A.: Existence and uniqueness of maximal regular flows for non-smooth vector fields. *Arch. Ration. Mech. Anal.* **218**(2), 1043–1081, 2015. https://doi.org/10.1007/s00205-015-0875-9

18. AUBIN, J.P.: Un théorème de compacité. *C. R. Acad. Sci. Paris* **256**, 5042–5044, 1963

19. BIANCHINI, S., BONICATTO, P.: A uniqueness result for the decomposition of vector fields in $\mathbb{R}^d$. *Invent. Math.* **220**(1), 255–393, 2020. https://doi.org/10.1007/s00222-019-00928-8

20. BOUCHUT, F.: Renormalized solutions to the Vlasov equation with coefficients of bounded variation. *Arch. Ration. Mech. Anal.* **157**(1), 75–90, 2001. https://doi.org/10.1007/PL00004237

21. BOUCHUT, F., DESVILLETTES, L.: On two-dimensional Hamiltonian transport equations with continuous coefficients. *Differ. Integ. Equ.* **14**(8), 1015–1024, 2001

22. BOUCHUT, F., JAMES, F.: One-dimensional transport equations with discontinuous coefficients. *Nonlinear Anal.* **32**(7), 891–933, 1998. https://doi.org/10.1016/S0362-546X(97)00536-1

23. BOUCHUT, F., JAMES, F.: Differentiability with respect to initial data for a scalar conservation law. In: Hyperbolic problems: theory, numerics, applications, Vol. I (Zürich, 1998), Internat. Ser. Numer. Math., vol 129. Birkhäuser, Basel, p 113–118 1999

24. BOUCHUT, F., JAMES, F.: Duality solutions for pressureless gases, monotone scalar conservation laws, and uniqueness. *Comm. Partial Differ. Equ.* **24**(11–12), 2173–2189, 1999. https://doi.org/10.1080/03605309908821498

25. BOUCHUT, F., JAMES, F., MANCINI, S.: Uniqueness and weak stability for multidimensional transport equations with one-sided Lipschitz coefficient. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **4**(1), 1–25, 2005

26. BRESSAN, A.: An ill posed Cauchy problem for a hyperbolic system in two space dimensions. *Rend. Sem. Mat. Univ. Padova* **110**, 103–117, 2003

27. BRESSAN, A.: A lemma and a conjecture on the cost of rearrangements. *Rend. Sem. Mat. Univ. Padova* **110**, 97–102, 2003

28. CAPUZZO DOLCETTA, I., PERTHAME, B.: On some analogy between different approaches to first order PDE's with nonsmooth coefficients. *Adv. Math. Sci. Appl.* **6**(2), 689–703, 1996

29. CARDALIAGUET, P., SOUGANIDIS, P.E.: On first order mean field game systems with a common noise. *Ann. Appl. Probab.* **32**(3), 2289–2326, 2022. https://doi.org/10.1214/21-aap1734

30. CARDALIAGUET, P., SEEGER, B., SOUGANIDIS, P.: Mean field games with common noise and degenerate idiosyncratic noise. 2023. Preprint at https://arxiv.org/abs/2207.10209

31. CHAMPAGNAT, N., JABIN, P.E.: Well posedness in any dimension for Hamiltonian flows with non BV force terms. *Comm. Partial Differ. Equ.* **35**(5), 786–816, 2010. https://doi.org/10.1080/03605301003646705

32. CHAMPAGNAT, N., JABIN, P.E.: Strong solutions to stochastic differential equations with rough coefficients. *Ann. Probab.* **46**(3), 1498–1541, 2018. https://doi.org/10.1214/17-AOP1208

33. CONWAY, E.D.: Generalized solutions of linear differential equations with discontinuous coefficients and the uniqueness question for multidimensional quasilinear conservation laws. *J. Math. Anal. Appl.* **18**, 238–251, 1967. https://doi.org/10.1016/0022-247X(67)90054-6

34. CRANDALL, M.G., ISHII, H.: The maximum principle for semicontinuous functions. *Differ. Integ. Equ.* **3**(6), 1001–1014, 1990

35. CRANDALL, M.G., ISHII, H., LIONS, P.L.: User's guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc.* **27**(1), 1–67, 1992. https://doi.org/10.1090/S0273-0979-1992-00266-5

36. CRIPPA, G., DE LELLIS, C.: Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.* **616**, 15–46, 2008. https://doi.org/10.1515/CRELLE.2008.016

37. CULLEN, M., FELDMAN, M.: Lagrangian solutions of semigeostrophic equations in physical space. *SIAM J. Math. Anal.* **37**(5), 1371–1395, 2006. https://doi.org/10.1137/040615444

38. DAFERMOS, C.M.: Generalized characteristics and the structure of solutions of hyperbolic conservation laws. *Indiana Univ. Math. J.* **26**(6), 1097–1119, 1977. https://doi.org/10.1512/iumj.1977.26.26088

39. DAVIE, A.M.: Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not IMRN*, 2007. https://doi.org/10.1093/imrn/rnm124

40. DIPERNA, R.J., LIONS, P.L.: Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98**(3), 511–547, 1989. https://doi.org/10.1007/BF01393835

41. FIGALLI, A.: Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.* **254**(1), 109–153, 2008. https://doi.org/10.1016/j.jfa.2007.09.020

42. FILIPPOV, A.F.: Differential equations with discontinuous right-hand side. *Mat. Sb* **51**(93), 99–128, 1960

43. GODLEWSKI, E., OLAZABAL, M., RAVIART, P.A.: On the linearization of hyperbolic systems of conservation laws. Application to stability. In: Équations aux dérivées partielles et applications. Gauthier-Villars, Éd. Sci. Méd. Elsevier, Paris, p 549–570 1998

44. GODLEWSKI, E., OLAZABAL, M., RAVIART, P.A.: On the linearization of systems of conservation laws for fluids at a material contact discontinuity. *J. Math. Pures. Appl.* **78**(10), 1013–1042, 1999. https://doi.org/10.1016/S0021-7824(99)00136-1

45. HAURAY, M.: On two-dimensional Hamiltonian transport equations with L_{loc}^p coefficients. *Ann. Inst. H Poincaré C Anal. Non Linéaire* **20**(4), 625–644, 2003. https://doi.org/10.1016/S0294-1449(02)00015-X

46. ISHII, H.: Hamilton-Jacobi equations with discontinuous Hamiltonians on arbitrary open sets. *Bull. Fac. Sci. Engrg. Chuo Univ.* **28**, 33–77, 1985

47. KARIMGHASEMI, S., MÜLLER, S., WESTDICKENBERG, M.: Flow solutions of transport equations. *Comm. Partial Differ. Equ.* **46**(1), 98–134, 2021. https://doi.org/10.1080/03605302.2020.1831018

48. KEYFITZ, B.L., KRANZER, H.C.: A system of nonstrictly hyperbolic conservation laws arising in elasticity theory. *Arch. Rational Mech. Anal.* **72**(3), 219–241, 1979. https://doi.org/10.1007/BF00281590

49. LASRY, J.M., LIONS, P.L.: Mean field games. *Jpn. J. Math.* **2**(1), 229–260, 2007. https://doi.org/10.1007/s11537-007-0657-8

50. LE BRIS, C., LIONS, P.L.: Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications. *Ann. Mat. Pura Appl.* **183**(1), 97–130, 2004. https://doi.org/10.1007/s10231-003-0082-4

51. LE BRIS, C., LIONS, P.L.: Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients. *Comm. Partial Differ. Equ.* **33**(7–9), 1272–1317, 2008. https://doi.org/10.1080/03605300801970952

52. LE BRIS, C., LIONS, P.L.: ([2019] © 2019) Parabolic equations with irregular data and related issues—applications to stochastic differential equations, De Gruyter Series in Applied and Numerical Mathematics, vol 4. De Gruyter, Berlin

53. LEFLOCH, P.: An existence and uniqueness result for two nonstrictly hyperbolic systems. In: Nonlinear evolution equations that change type, IMA Vol. Math. Appl., vol 27. Springer, New York, p 126–138, 1990. https://doi.org/10.1007/978-1-4613-9049-7_10

54. LIONS, J.L.: Quelques méthodes de résolution des problèmes aux limites non linéaires. Gauthier-Villars, Paris, Dunod, Paris (1969)

55. LIONS, P.L., PERTHAME, B.: Remarks on Hamilton-Jacobi equations with measurable time-dependent Hamiltonians. *Nonlinear Anal.* **11**(5), 613–621, 1987. https://doi.org/10.1016/0362-546X(87)90076-9

56. LIONS, P.L., SEEGER, B.: Linear and nonlinear transport equations with coordinate-wise increasing velocity fields. Inst H Poincaré C Anal Non Linéaire, 2024. https://doi.org/10.4171/AIHPC/133, published online first

57. NUNZIANTE, D.: Uniqueness of viscosity solutions of fully nonlinear second order parabolic equations with discontinuous time-dependence. *Differ. Integral Equ.* **3**(1), 77–91, 1990

58. NUNZIANTE, D.: Existence and uniqueness of unbounded viscosity solutions of parabolic equations with discontinuous time-dependence. *Nonlinear Anal.* **18**(11), 1033–1062, 1992. https://doi.org/10.1016/0362-546X(92)90194-J
59. PETROVA, G., POPOV, B.: Linear transport equations with discontinuous coefficients. *Comm. Partial Differ. Equ.* **24**(9–10), 1849–1873, 1999. https://doi.org/10.1080/03605309908821484
60. PETROVA, G., POPOV, B.: Linear transport equations with μ-monotone coefficients. *J. Math. Anal. Appl.* **260**(2), 307–324, 2001. https://doi.org/10.1006/jmaa.2000.7427
61. POUPAUD, F., RASCLE, M.: Measure solutions to the linear multi-dimensional transport equation with non-smooth coefficients. *Comm. Partial Differ. Equ.* **22**(1–2), 337–358, 1997. https://doi.org/10.1080/03605309708821265
62. RÖCKNER, M., ZHANG, X.: Weak uniqueness of Fokker-Planck equations with degenerate and bounded coefficients. *C. R. Math. Acad. Sci. Paris* **348**(7–8), 435–438, 2010. https://doi.org/10.1016/j.crma.2010.01.001
63. SUQUET, P.M.: Existence et régularité des solutions des équations de la plasticité. *C. R. Acad. Sci. Paris Sér A-B* **286**(24), A1201–A1204, 1978
64. TREVISAN, D.: Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron J. Probab.* **21**, 41, 2016. https://doi.org/10.1214/16-EJP4453
65. VERETENNIKOV, A.J.: Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb (NS)* **111**, 434–452, 1980
66. YAMADA, T., WATANABE, S.: On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11**, 155–167, 1971. https://doi.org/10.1215/kjm/1250523691
67. ZHENG, Y.X., MAJDA, A.: Existence of global weak solutions to one-component Vlasov-Poisson and Fokker-Planck-Poisson systems in one space dimension with measures as initial data. *Comm. Pure Appl Math.* **47**(10), 1365–1401, 1994. https://doi.org/10.1002/cpa.3160471004

P.-L. Lions
Université Paris-Dauphine and Collège de France,
Pl. du Maréchal de Lattre de Tassigny,
75016 Paris
France.
e-mail: lions@ceremade.dauphine.fr

and

B. Seeger
University of North Carolina at Chapel Hill,
318 Hanes Hall, CB #3260,
Chapel Hill
NC
27599 USA.
e-mail: bseeger@unc.edu