

Linear and nonlinear transport equations with coordinate-wise increasing velocity fields

Pierre-Louis Lions and Benjamin Seeger

Abstract. We consider linear and nonlinear transport equations with irregular velocity fields, motivated by models coming from mean field games. The velocity fields are assumed to increase in each coordinate, and the divergence therefore fails to be absolutely continuous with respect to the Lebesgue measure in general. For such velocity fields, the well-posedness of first- and second-order linear transport equations in Lebesgue spaces is established, as well as the existence and uniqueness of regular ODE and SDE Lagrangian flows. These results are then applied to the study of certain nonconservative, nonlinear systems of transport type, which are used to model mean field games in a finite state space. A notion of weak solution is identified for which unique minimal and maximal solutions exist, which do not coincide in general. A selection-by-noise result is established for a relevant example to demonstrate that different types of noise can select any of the admissible solutions in the vanishing noise limit.

Contents

1. Introduction	971
2. Preliminary results	979
3. The regular Lagrangian flow	986
4. Linear transport equations with increasing drift	1001
5. Nonlinear transport systems	1017
References	1033

1. Introduction

This paper has two main purposes. First we develop a well-posedness theory for first- and second-order linear transport equations for a new class of irregular velocity fields, as well as the corresponding ODE and SDE regular Lagrangian flows. We then apply the results to the study of certain nonlinear transport systems, motivated in particular by applications to mean field games on a finite state space.

Mathematics Subject Classification 2020: 35F50 (primary); 35F10, 34A36, 60H10, 91A16 (secondary).
Keywords: transport equations, regular Lagrangian flow, Fokker–Planck equation, SDEs, mean field games, stochastic selection.

For a fixed, finite time horizon $T > 0$, we study both the terminal value problem (TVP) for the nonconservative equation

$$-\partial_t u - b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_0, \quad (1.1)$$

as well as the dual, initial value problem (IVP) for the conservative equation

$$\partial_t f + \operatorname{div}(b(t, x)f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad f(0, \cdot) = f_0, \quad (1.2)$$

under the assumption that b is coordinate-by-coordinate (semi)-increasing:

$$\begin{aligned} \partial_{x_j} b^i(t, \cdot) &\geq -C(t) \delta_{ij} \quad \text{for all } t \in [0, T], i, j = 1, 2, \dots, d, \\ \text{and some } C &\in L^1_+[0, T]. \end{aligned} \quad (1.3)$$

The divergence of such a vector field is bounded from below, which means that, formally, the flow defined by

$$\partial_t \phi_{t,s}(x) = b(t, \phi_{t,s}(x)), \quad t \in [s, T], \quad \phi_{s,s}(x) = x \quad (1.4)$$

will not concentrate at null sets. This indicates that the two problems (1.1) and (1.2) are amenable to a solution theory in Lebesgue spaces.

On the other hand, the measure $\operatorname{div} b$ is not in general absolutely continuous with respect to Lebesgue measure, and this leads to the formation of vacuum for $t > s$. It is therefore the case that the existing theory of renormalized solutions, initiated by DiPerna and the first author [30] for Sobolev velocity fields and extended to the case $b \in \operatorname{BV}_{\operatorname{loc}}$ and $\operatorname{div} b \in L^\infty$ by Ambrosio [2], does not apply to our present situation. In particular, the two problems (1.1) and (1.2) cannot be covered with a unified theory, due to the fact that (1.1) cannot be understood in the distributional sense if $\operatorname{div} b$ is not absolutely continuous with respect to Lebesgue measure and $u \in L^1_{\operatorname{loc}}$. Nevertheless, we exploit the dual relationship between the two problems, and provide a link to the forward, regular Lagrangian flow (1.4). Analogous results are also proved for the degenerate, second-order equations

$$-\partial_t u - b(t, x) \cdot \nabla u - \operatorname{tr}[a(t, x) \nabla^2 u] = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T \quad (1.5)$$

and

$$\partial_t f + \operatorname{div}(b(t, x)f) - \nabla^2 \cdot (a(t, x)f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad f(0, \cdot) = f_0, \quad (1.6)$$

where b satisfies (1.3), $a(t, x) = \frac{1}{2} \sigma(t, x) \sigma(t, x)^\top$ is a nonnegative, symmetric matrix, and $\sigma \in \mathbb{R}^{d \times m}$; and the SDE flow

$$d_t \Phi_{t,s}(x) = b(t, \Phi_{t,s}(x)) dt + \sigma(t, \Phi_{t,s}(x)) dW_t, \quad (1.7)$$

where W is an m -dimensional Brownian motion.

We then turn to the study of nonlinear, nonconservative systems of transport type that take the form

$$-\partial_t u - f(t, x, u) \cdot \nabla u = g(t, x, u) \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T. \quad (1.8)$$

Here, u , f , and g are vector valued, with¹ $f \in \mathbb{R}^d$ and $g, u \in \mathbb{R}^m$ for some integers $m, d \geq 1$, f and g are local functions of $(t, x, u) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^m$, and the equation reads, for each coordinate $i = 1, 2, \dots, m$,

$$-\partial_t u^i - \sum_{j=1}^d f^j(t, x, u) \partial_{x_j} u^i = g^i(t, x, u).$$

The primary motivation for the consideration of (1.8) comes from the study of mean field games (MFG). These are models for large populations of interacting rational agents, which strategize in order to optimize an outcome, based on the collective behavior of the remaining population, while subject to environmental influences. The master equation for mean field games with a (finite) discrete state space takes the general form of system (1.8) with $d = m$, as described in [47]; see also [10, 35]. Alternatively, systems of the form (1.8) arise upon exploiting dimension reduction techniques for continuum-state MFG models in which the various data depend on the probability distribution of players through a finite number of observables, that is,

$$\mu \mapsto u\left(t, \int \Phi d\mu\right)$$

for probability measures μ and some given continuous \mathbb{R}^d -valued function Φ . This connection is explored by the authors and Lasry in [44]. We note also that the special case where $d = m$, $f(t, x, u) = -u$, and $g(t, x, u) = 0$ leads to the system

$$-\partial_t u + u \cdot \nabla u = 0, \quad (1.9)$$

which arises in certain models describing the flow of compressible gasses at low density with negligible pressure [38, 49].

The nonlinear equation (1.8) can formally be connected to a system of characteristic ODEs on $\mathbb{R}^d \times \mathbb{R}^m$. In order to draw the analogy to MFG PDE systems and the master equation in a continuum state space (see for instance [17]), it is convenient to represent the characteristics as the forward backward system

$$\begin{cases} -\partial_s U_{s,t}(x) = g(s, X_{s,t}(x), U_{s,t}(x)), & U_{T,t}(x) = u_T(X_{T,t}(x)), \\ \partial_s X_{s,t}(x) = f(s, X_{s,t}(x), U_{s,t}(x)), & X_{t,t}(x) = x. \end{cases} \quad (1.10)$$

¹Note that the nonlinearity f in equation (1.8) is different from the solutions f of the conservative equations (1.2) or (1.6).

If f , g , and u_T are smooth and the interval $[t, T]$ is sufficiently small, then (1.10) can be uniquely solved, and the unique, smooth solution u of (1.8) is given by $u(t, x) := U_{t,t}(x)$. The argument fails for arbitrarily long time intervals, in view of the coupling between X and the terminal condition for U .²

The monotone regime is explored in [10, 47], that is, $d = m$ and $(-g, f): \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ and $u_T: \mathbb{R}^d \rightarrow \mathbb{R}^d$ are smooth and monotone (one of which is strictly monotone). In that case, (1.10), and therefore (1.8), can be uniquely solved on any time interval. This regime is exactly analogous to the monotonicity condition of Lasry and the first author for MFG systems with continuum state space [41–43], and, as in that setting, strong regularity and stability results can be established for (1.8). The monotone regime also allows for a well-posed notion of global weak continuous solutions of (1.8), even when f , g , and u_T fail to be smooth [8].

If there exist functions $H: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ and $v_T: \mathbb{R}^d \rightarrow \mathbb{R}$ such that $(g, f) = (\nabla_x, \nabla_p)H$ and $u_T = \nabla v_T$ (the so-called potential regime), then, formally, one expects $u(t, \cdot) = \nabla v(t, \cdot)$, where v solves the Hamilton–Jacobi equation

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

for which global, weak continuous solutions can be understood with the theory of viscosity solutions [24]. Possibly discontinuous weak solutions of (1.8) can then be indirectly understood as the distributional derivative of v , an approach which is taken in [21]. In the special case where $d = m = 1$, (1.8) can be studied with the theory of entropy solutions of scalar conservation laws [40].

In this paper, using the theory developed here for linear equations, we identify new regimes of assumptions on f , g , and u_T for which the notion of a weak, possibly discontinuous, solution can be identified for any dimensions $d, m \geq 1$. Under a certain ordering structure, the existence of unique maximal and minimal solutions, which do not coincide in general, is established. This nonuniqueness is further explored from the viewpoint of stochastic selection, and we prove, for a specific but informative example, that *any* of the family of solutions can be distinguished by a certain vanishing noise limit, indicating that the choice of a mean field game equilibrium is very sensitive to the manner in which low-level, systemic noise is introduced into the model.

1.1. Summary of main results

We list the main results of the paper here, in an informal setting. More precise statements and discussions can be found within the body of the paper.

The (semi)-increasing condition (1.3) implies that $b \in \text{BV}_{\text{loc}}$. We emphasize again, however, that the measure $\text{div } b$ will in general have a singular part with respect to

²When written in forward form, the breakdown manifests as a failure to invert the characteristics describing the state variable $x \in \mathbb{R}^d$, which may cross in finite time.

Lebesgue measure, and we therefore cannot appeal to the existing results on renormalized solutions to transport equations with irregular velocity fields. We do not give a full account of the vast literature for such problems, but refer the reader to the thorough surveys [3–6] and the references therein.

Our approach to the first-order transport problem is to study the well-posedness of the regular Lagrangian flow for (1.4) directly, rather than using PDE methods. The assumption (1.3) on b allows for a comparison principle with respect to the partial order

$$x, y \in \mathbb{R}^d, \quad x \leq y \Leftrightarrow x_i \leq y_i \text{ for all } i = 1, 2, \dots, d. \quad (1.11)$$

A careful regularization procedure then leads to the existence of minimal and maximal flows, which coincide a.e., and we have the following result (see Section 3 below for more precise statements):

Theorem 1.1. *Assume b satisfies (1.3). Then, for a.e. $x \in \mathbb{R}^d$, there exists a unique absolutely continuous solution of (1.4), and there exists a constant $C > 0$ such that, for all $0 \leq s \leq t \leq T$,*

$$|\phi_{t,s}^{-1}(A)| \leq C|A| \quad \text{for all measurable } A \subset \mathbb{R}^d.$$

If $(b^\varepsilon)_{\varepsilon>0}$ is a family of smooth approximations of b and $(\phi^\varepsilon)_{\varepsilon>0}$ are the corresponding flows, then, as $\varepsilon \rightarrow 0$, $\phi^\varepsilon \rightarrow \phi$ in $C_t L_{x,\text{loc}}^p$ and $L_{x,\text{loc}}^p W_t^{1,1}$ for all $p \in [1, \infty)$.

We also obtain analogous results for the SDE (1.7). Degenerate linear parabolic equations and SDEs with irregular data have been studied in a number of works that generalize the DiPerna–Lions theory and the Ambrosio superposition principle; see [22, 31, 45, 46, 58]. A common source of difficulty involves the dependence of σ on the spatial variable, even if it is smooth. This is the case, for instance, when $b \in \text{BV}_{\text{loc}}$ and $\text{div } b \in L^\infty$ treated by Figalli [31], or when b satisfies a one-sided Lipschitz condition from below, as considered by the authors in [48]; in both settings, the results are constrained to $\sigma(t, x) = \sigma(t)$ constant in \mathbb{R}^d . In our present setting, we can relax the spatial dependence, and we assume that σ is Lipschitz and satisfies

$$\begin{aligned} \sigma^{ik}(t, x) &= \sigma^{ik}(t, x_i) \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d, \\ i &= 1, 2, \dots, d, \quad k = 1, \dots, m. \end{aligned}$$

Then, in Section 4, we turn to the study of the linear transport equations (1.1)–(1.2), as well as the second-order equations (1.5)–(1.6), which can be related to the ODE and SDE flows in Section 3.

Theorem 1.2. *Assume b satisfies (1.3). Then the flow (1.4) (resp. (1.7)) gives rise to continuous solution operators on L_{loc}^p (resp. L^p), $p \in (1, \infty)$ for the dual problems (1.1)–(1.2) (resp. (1.5)–(1.6)). The resulting solutions are stable under regularizations of b or vanishing viscosity limits in the spaces $C_t L_x^p$, with the convergence being strong for the nonconservative equations (1.1)/(1.5) and weak for the conservative equations (1.2)/(1.6).*

The solution operator for the nonconservative equation (1.1) cannot be made sense of in the sense of distributions, because the measure $\operatorname{div} b$ can have a singular part. We nevertheless provide a PDE-based characterization for solutions that are *increasing or decreasing* with respect to the partial order (1.11). In order to give meaning to the ill-defined product $b \cdot \nabla u$ in this context, we introduce mollifications of a one-sided nature that lead to commutator errors which are shown to possess a sign; this is to be compared with the renormalization theory initiated in [30], in which the commutator errors are shown to converge to zero with the convolution parameter. This leads to the notions of sub- and supersolution for (1.1), which are proved to satisfy a comparison principle. The solution operator for (1.1) can thus be alternatively characterized in terms of these regularizations, and, moreover, the regular Lagrangian flow (1.4) can be recovered as the (vector)-valued solution of the terminal value problem (1.1) with $u_T(x) = x$. These two viewpoints on the transport equation and ODE flow are instrumental in our understanding of the nonlinear equations to follow.

The continuity equation (1.2) (or the Fokker–Planck equation (1.6) in the second-order case) can then be related to the nonconservative equation through duality. Importantly, arbitrary distributional solutions of (1.2) are not unique in general. We prove that, if $f_0 \geq 0$, then there exist unique distributional solutions of (1.2) and (1.6) which coincide with the duality solution. Moreover, this result is proved independently of the superposition principle; instead, we use the duality with the nonconservative equation, and the characterization of its solutions in terms of one-sided regularizations.

A consequence of the uniqueness of nonnegative distributional solutions of (1.2) is that, if f and $|f|$ satisfy (1.2) in the sense of distributions, then f is the “good” (duality) solution (see Corollary 4.1 below). We do not know whether this property characterizes the duality solution, or, in other words, whether the duality solution satisfies the renormalization property in general. This should be compared with [48], where the authors resolve the same questions for half-Lipschitz velocity fields.

The paper concludes in Section 5 with the study of the nonlinear equation (1.8) and the associated system (1.10). We operate under the assumption that the discontinuous nonlinearities f and g satisfy, for some $C \in L^1_+[0, T]$,

$$\left\{ \begin{array}{l} \text{for all } i, j = 1, 2, \dots, d, \ k, \ell = 1, 2, \dots, m, \\ \text{and a.e. } t \in [0, T], \text{ on } \mathbb{R}^d \times \mathbb{R}^m, \\ \partial_{x_i} f^j(t, \cdot, \cdot) \geq -C(t) \delta_{ij}, \ \partial_{u_k} g^\ell(t, \cdot, \cdot) \geq -C(t) \delta_{k\ell}, \\ \partial_{x_i} g^\ell \leq 0, \text{ and } \partial_{u_k} f^j \leq 0. \end{array} \right. \quad (1.12)$$

Observe that (1.12) is satisfied with $C \equiv 0$ by the particular example of the Burgers-like equation (1.9).

We develop a theory for solutions of (1.8) that are *decreasing* with respect to the partial order (1.11). The first observation is that the decreasing property is propagated, formally, by the solution operator. On the other hand, shocks form in finite time, and so, even if u_T , f , and g are smooth, $u(t, \cdot)$ will develop discontinuities for some $t < T$ in general,

requiring a notion of weak solution. We next note that, under the above assumptions, if u is decreasing, then the velocity field $b(t, x) := f(t, x, u(t, x))$ satisfies (1.3). Solutions of the nonlinear equation (1.8) can then be understood as fixed points for the linear problem (1.1), and at the same time through the system of forward-backward characteristics (1.10), using the theory for the regular Lagrangian flows in Section 3.

Theorem 1.3. *Assume f and g satisfy (1.12) and u_T is decreasing. Then there exist maximal and minimal decreasing solutions u^+ and u^- of (1.8) in the fixed-point sense. If u is any other such solution, then $u^- \leq u \leq u^+$.*

Continuous decreasing solutions of (1.8) satisfy a comparison principle, which in particular implies that u^+ is below every continuous supersolution and u^- is above every continuous subsolution. In general, u^- and u^+ do not coincide, which must then be a consequence of the formation of discontinuities. This nonuniqueness of solutions is closely related to the question of multiplying distributions of limited regularity. Indeed, in equation (1.8), the product $f(t, x, u) \cdot \nabla u$ cannot be defined in a stable (with respect to regularizations) way in view of the fact that, in general, $u \in \text{BV}_{\text{loc}}$ and ∇u is a locally finite measure.

The well-posedness of both strong and weak solutions of MFG master equations in the continuum state space setting has been explored under various sets of monotonicity assumptions [1, 9, 11, 17, 18, 20, 33, 34, 36, 37, 50]. The approach in our setting, which involves appealing to Tarski's fixed-point theorem for increasing functions on lattices [57], has also been taken in the continuum state space setting, where maximal and minimal solutions were found under related assumptions; see for instance [27–29, 51]. The partial order used in [51] comes from the notion of stochastic dominance for probability measures. We note that, for equation (1.8) posed on an infinite-dimensional Hilbert space of L^2 random variables, the partial order (1.11) is related to the analogous notion of stochastic dominance for random variables, and we aim to study the infinite-dimensional version of (1.8) in future work.

We explore the nonuniqueness issue in more detail for the Burgers equation (1.9) in one dimension, where the decreasing terminal value has a single discontinuity at 0:

$$-\partial_t u + u \partial_x u = 0 \quad \text{in } (0, T) \times \mathbb{R}, \quad u(T, x) = \mathbf{1}\{x \leq 0\}. \quad (1.13)$$

It turns out that (1.13) admits infinitely many fixed-point solutions, consisting of a shock traveling with variable speed between 0 and 1. Of course, (1.13) can be reframed as a scalar conservation law, whose unique entropy solution is the shock-wave solution with speed $1/2$. We note that the notion of entropy solution does not extend to the nonconservative equations (1.8) or (1.9).

We characterize the family of fixed-point solutions (1.13) as limits under distinct types of regularizations of equation (1.13).

Theorem 1.4. *For any $c \in W^{2,1}([0, T])$ satisfying $c(T) = 0$ and $-c' \in (0, 1)$, there exists $\theta_c \in L^1([0, T])$ such that, if u_T^ε is smooth, $u_T^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \mathbf{1}_{(-\infty, 0)}$ in L^1_{loc} , and u^ε is the unique*

classical solution of

$$-\partial_t u^\varepsilon + u^\varepsilon \partial_x u^\varepsilon = \varepsilon(\partial_x^2 u^\varepsilon + \theta_c(t)|\partial_x u^\varepsilon|^2), \quad u^\varepsilon(T, \cdot) = u_T^\varepsilon, \quad (1.14)$$

then, for all $1 \leq p < \infty$, as $\varepsilon \rightarrow 0$, $u^\varepsilon \rightarrow \mathbf{1}\{x \leq c(t)\}$ strongly in $C([0, T], L_{\text{loc}}^p)$.

We interpret Theorem 1.4 as a selection-by-noise result for the nonunique problem (1.13). Indeed, the result can be reformulated on the level of system (1.10), which, for (1.14), becomes the forward-backward system of SDEs

$$\begin{cases} -d_s U_{s,t}^\varepsilon(x) = Z_{s,t}^\varepsilon(x) dW_s - \frac{1}{2} \theta_c(s) Z_{s,t}^\varepsilon(x)^2 ds, & U_{T,t}^\varepsilon(x) = u_T^\varepsilon(X_{T,t}^\varepsilon(x)), \\ d_s X_{s,t}^\varepsilon(x) = -U_{s,t}^\varepsilon(x) ds + \sqrt{2\varepsilon} dW_s, & X_{t,t}^\varepsilon(x) = x. \end{cases}$$

Various selection methods have been proposed to study mean field game models that do not admit unique solutions, and we refer in particular to [21, 26] for problems involving stochastic selection. Our result is distinguished by the consideration of several different descriptions of small noise, each of which selects a different solution for the deterministic problem in the vanishing noise limit.

1.2. A note on velocity fields with a one-sided Lipschitz condition

Let us remark that the regime in which b satisfies (1.3) shares many similarities with the setting in which b is half-Lipschitz from below, that is,

$$(b(t, x) - b(t, y)) \cdot (x - y) \geq -C(t)|x - y|^2 \quad \text{for } (t, x, y) \in [0, T] \times \mathbb{R}^{2d} \\ \text{and some } C \in L_+^1([0, T]). \quad (1.15)$$

A key commonality in both settings is that $\text{div } b$ is bounded from below, but not necessarily absolutely continuous with respect to Lebesgue measure. Transport equations and flows for velocity fields satisfying (1.15) have been studied from a variety of different viewpoints [12, 13, 19, 23, 53–55], and in [48], the authors obtain very similar results to those described above regarding the existence, uniqueness, and stability of the regular Lagrangian flow forward in time, as well as proving well-posedness and studying properties and characterizations of solutions to problems (1.1) and (1.2) in Lebesgue spaces.

A key difference between the two regimes is the behavior of the flow for (1.4) in the compressive direction, that is, backward in time. For velocity fields satisfying the half-Lipschitz condition (1.15), the backward ODE is uniquely solvable for all $x \in \mathbb{R}^d$. Moreover, the resulting backward flow is Lipschitz continuous, and it can be identified as the left-inverse to the forward, regular Lagrangian flow; see [48] for more precise statements, as well as new characterizations of the time-reversed versions of (1.1) and (1.2).

On the other hand, when b satisfies (1.3), the backward problem (1.4) is not in general unique for every $x \in \mathbb{R}^d$, nor is it true that a globally Lipschitz flow can always be found. We note that, even for examples where the backward flow has a unique solution

for Lebesgue-a.e. $x \in \mathbb{R}^d$, neither the stability nor the solvability of the time-reversed versions of (1.1)–(1.2) in Lebesgue spaces can be expected to hold, because, in general, any backward flow solution to (1.4) will concentrate on sets of Lebesgue-measure zero. For a detailed discussion and examples, see Sections 3.4 and 4.3 below.

We remark that, when $d = 1$, assumption (1.3) coincides exactly with the one-sided Lipschitz bound, and so, in the one-dimensional setting, this paper presents a counterpart to the approaches taken in [12, 13, 48]. In particular, in those works, the analysis is based primarily on the properties of the (well-defined) backward flow and the adjoint conservative continuity equation, whereas here we begin with establishing the well-posedness of the forward Lagrangian flow, an approach which generalizes well to multiple dimensions under assumption (1.3).

1.3. Notation

Given a bounded function $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$, ϕ_* and ϕ^* denote the lower- and upper-semi-continuous envelopes, and, if ϕ is \mathbb{R}^m -valued, the same notation is used coordinate by coordinate.

We often denote arbitrary functions spaces on \mathbb{R}^d as $X(\mathbb{R}^d) = X$ when there is no ambiguity over the domain. We denote the space of bounded, Lipschitz continuous functions by $C^{0,1}$. For $p \in [1, \infty]$, L_w^p and $L_{w,*}^p$ denote the spaces of p -integrable functions with respectively the weak and weak- \star topologies. The space of locally p -integrable functions with the topology of local L^p -convergence is denoted by L_{loc}^p , and $L_{\text{loc},w}^p$ and $L_{\text{loc},w,*}^p$ are understood accordingly.

The notation $\mathbf{1}$ denotes the vector $(1, 1, \dots, 1)$ in Euclidean space, the dimension being clear from the context. Given two sets A and B , $A \triangle B := (A \setminus B) \cup (B \setminus A)$.

2. Preliminary results

This section contains a collection of results regarding vector-valued notions of increasing/decreasing, as well as a vector-valued maximum principle.

2.1. Properties of increasing functions

We first introduce a partial order on \mathbb{R}^d that is used throughout the paper.

Definition 2.1. For $x, y \in \mathbb{R}^d$, we will write

$$x \leq y \quad \text{if } x_i \leq y_i \text{ for all } i = 1, 2, \dots, d. \quad (2.1)$$

Given $a, b \in \mathbb{R}^d$, $a \leq b$, we denote by $[a, b]$ the cube $\prod_{i=1}^d [a_i, b_i]$. A function $\phi: \mathbb{R}^d \rightarrow \mathbb{R}^m$, $d, m \in \mathbb{N}$, is said to be increasing if $\phi(x) \leq \phi(y)$ whenever $x \leq y$. Equivalently, ϕ is increasing if, for each $i = 1, 2, \dots, m$, ϕ^i is increasing in the x_j -coordinate for each $j = 1, 2, \dots, d$.

Lemma 2.1. Assume that $\phi: \mathbb{R}^d \rightarrow \mathbb{R}^m$ is increasing. Then $\phi \in \text{BV}_{\text{loc}}(\mathbb{R}^d)$, and, for each $i = 1, 2, \dots, m$,

$$\liminf_{y \rightarrow x} \phi^i(y) = \liminf_{y \rightarrow x, y \not\geq x} \phi^i(y) \quad \text{and} \quad \limsup_{y \rightarrow x} \phi^i(y) = \limsup_{y \rightarrow x, y \not\leq x} \phi^i(y).$$

Proof. For all $j = 1, 2, \dots, d$ and $i = 1, 2, \dots, m$, the distribution $\partial_j \phi^i$ is nonnegative, and is therefore a locally finite measure.

For any sequence $y_n \rightarrow x$, if $y_n \not\geq x$, then y_n may be replaced with a value y'_n such that $y'_n \leq x$, $y'_n \leq y_n$, $y'_n \neq x$, $\phi^i(y'_n) \leq \phi^i(x)$, and $\lim_{n \rightarrow \infty} y'_n = x$. A similar statement holds if $y_n \not\leq x$ for all n , and the claim follows. ■

Lemma 2.1 implies that each component of an increasing function $\phi: \mathbb{R}^d \rightarrow \mathbb{R}^m$ has limits from both the “left” and “right”. We will call an increasing function ϕ càdlàg if

$$\phi^i(x) = \limsup_{y \rightarrow x} \phi^i(y) = \limsup_{y \rightarrow x, y \not\geq x} \phi^i(y) \quad \text{for } i = 1, 2, \dots, m. \quad (2.2)$$

Remark 2.1. Given a nonnegative measure μ , the repartition function

$$\phi(x) := \int_{(-\infty, x_1)} \int_{(-\infty, x_2)} \cdots \int_{(-\infty, x_d)} d\mu$$

is an example of a càdlàg increasing function, but such functions do not cover the full range of increasing functions if $d \geq 2$; indeed, they are distinguished by the fact that mixed derivatives $\partial_{x_{\ell_1}} \partial_{x_{\ell_2}} \cdots \partial_{x_{\ell_k}} \phi$ for any distinct set $(\ell_j)_{j=1}^k \subset \{1, 2, \dots, d\}$ are measures even if $k \geq 2$.

Consider a smooth surface $\Gamma \subset \mathbb{R}^d$ that partitions \mathbb{R}^d into two open sets, that is, $\mathbb{R}^d = U_- \cup \Gamma \cup U_+$, $\Gamma = \partial U_+ = \partial U_-$, $U_+ \cap U_- = \emptyset$. Let n be the normal vector to Γ that, at all points of Γ , points inward to U_+ . If $n \geq 0$ everywhere on Γ , in the sense of (2.1), then $\phi = \mathbf{1}_{U_+ \cup \Gamma}$ is a càdlàg increasing function that is not a repartition function.

2.2. ABV functions

In one dimension, functions of bounded variation can be written as a difference of nondecreasing functions. With respect to the partial order (2.1), the generalization of this notion is a strict subspace of BV.

Definition 2.2. Given $-\infty < a_i \leq b_i < \infty$ for $i = 1, 2, \dots, d$ and a function

$$\phi: Q := \prod_{i=1}^d [a_i, b_i] \rightarrow \mathbb{R},$$

we say $\phi \in \text{ABV}(Q)$ if

$$\|\phi\|_{\text{ABV}(Q)} := \sup_{\gamma} \|\phi \circ \gamma\|_{\text{BV}([0,1])} < \infty,$$

where the supremum is taken over all curves $\gamma: [0, 1]^d \rightarrow Q$ such that $\gamma_i: [0, 1] \rightarrow [a_i, b_i]$ is increasing for all $i = 1, 2, \dots, d$. We will say $\phi \in \text{ABV} = \text{ABV}(\mathbb{R}^d)$ if $\|\phi\|_{\text{ABV}(Q)} < \infty$ for all boxes $Q \subset \mathbb{R}^d$.

For example, $C^{0,1} \subset \text{ABV}$. It is straightforward to see that $\text{ABV} = \text{BV}$ when $d = 1$.

Remark 2.2. Several generalizations of the notion of finite variation to multiple dimensions, besides the space BV , exist in the literature, and the one in Definition 2.2 is due to Arzelà [7]. It is a strictly smaller subspace than BV (indeed, unlike general BV functions, ABV functions are differentiable almost everywhere, as shown in Lemma 2.3 below). This notion of variation, along with several others, seems not to have had the same ubiquity in the theory of PDEs as the usual notion of BV , but is particularly relevant in this paper. More details about ABV functions, and many other notions of variation in multiple dimensions, can be found in [14].

Lemma 2.2. *Let $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$. Then $\phi \in \text{ABV}$ if and only if $\phi = \phi_1 - \phi_2$ for two increasing functions $\phi_1, \phi_2: \mathbb{R}^d \rightarrow \mathbb{R}$.*

Proof. Let $a, b \in \mathbb{R}^d$, $a \leq b$. If $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$ and $\gamma: [0, 1] \rightarrow [a, b]$ are increasing, then $\phi \circ \gamma$ is increasing, and thus $\|\phi \circ \gamma\|_{\text{BV}([0,1])} = \phi(b) - \phi(a)$. It follows that $\phi \in \text{ABV}$, and, by linearity, differences of increasing functions belong to ABV .

Now assume $\phi \in \text{ABV}([a, b])$, and, for $x \in [a, b]$, set

$$\phi_1(x) = \frac{1}{2}(\|\phi\|_{\text{ABV}([a,x])} + \phi(x)) \quad \text{and} \quad \phi_2(x) = \frac{1}{2}(\|\phi\|_{\text{ABV}([a,x])} - \phi(x)).$$

Then $\phi = \phi_1 - \phi_2$, and ϕ_1 and ϕ_2 are increasing. ■

Lemma 2.3. *If $\phi \in \text{ABV}$, then ϕ is almost everywhere continuous and differentiable.*

Proof. It suffices to prove the claim about differentiability, and, by Lemma 2.2, we may assume without loss of generality that ϕ is increasing.

We argue using the characterization by Stepanoff [56] of a.e.-differentiability, that is, we prove that, for almost every $x \in \mathbb{R}^d$,

$$\limsup_{y \rightarrow x} \frac{|\phi(y) - \phi(x)|}{|y - x|} < \infty,$$

or, equivalently,

$$\limsup_{y \rightarrow x} \frac{\phi(y) - \phi(x)}{|y - x|} < \infty \quad \text{and} \quad \liminf_{y \rightarrow x} \frac{\phi(y) - \phi(x)}{|y - x|} > -\infty.$$

Denote $Q_r := [0, r]^d$, and note that, because ϕ is increasing,

$$\limsup_{y \rightarrow x} \frac{\phi(y) - \phi(x)}{|y - x|} = \inf_{r > 0} \sup_{y \in x + Q_r} \frac{\phi(y) - \phi(x)}{|y - x|} = \inf_{r > 0} \frac{\phi(x + r\mathbf{1}) - \phi(x)}{\sqrt{d}r}.$$

The function $[0, \infty) \ni r \rightarrow \phi(x + r\mathbf{1})$ is increasing, and, thus, is differentiable almost everywhere in $[0, \infty)$. The finiteness of the above expression for almost every $x \in \mathbb{R}^d$ is then a consequence of Fubini's theorem. The argument for the \liminf is analogous. ■

Remark 2.3. In view of Lemmas 2.1 and 2.3, an increasing function is equal almost everywhere to a càdlàg function (that is, a function for which (2.2) holds), and we assume for the rest of the paper that increasing functions are càdlàg.

2.3. One-sided regularizations

It will be convenient at several times in the paper to specify regularizations of discontinuous functions that enjoy certain ordering properties. We describe two such regularization procedures here, each of which has merit in different situations.

We first discuss the inf- and sup-convolutions given, for some measurable function ϕ , by

$$C^\varepsilon[\phi](x) = \sup_{y \in \mathbb{R}^d} \left\{ \phi(x - y) - \frac{|y|}{\varepsilon} \right\} \quad \text{and} \quad C_\varepsilon[\phi](x) = \inf_{y \in \mathbb{R}^d} \left\{ \phi(x - y) + \frac{|y|}{\varepsilon} \right\}. \quad (2.3)$$

The following properties, which are either well known or easy to check, establish in particular the regularity of $C^\varepsilon[\phi]$ and of $C_\varepsilon[\phi]$ and their convergence to the appropriate semicontinuous envelopes ϕ^* and ϕ_* of ϕ .

Lemma 2.4. Assume that ϕ is measurable and, for some $M > 0$ and all $x \in \mathbb{R}^d$, $|\phi(x)| \leq M(1 + |x|)$, and let $C_\varepsilon[\phi]$ and $C^\varepsilon[\phi]$ be defined by (2.3), which are finite as long as $\varepsilon < M^{-1}$. Then,

- (a) for all $\varepsilon \in (0, M^{-1})$, $C_\varepsilon[\phi]$ and $C^\varepsilon[\phi]$ are Lipschitz with constant ε^{-1} , and $C_\varepsilon[\phi] \leq \phi \leq C^\varepsilon[\phi]$;
- (b) for all $x \in \mathbb{R}^d$, as $\varepsilon \rightarrow 0$, $C^\varepsilon[\phi](x) \searrow \phi^*(x)$ and $C_\varepsilon[\phi](x) \nearrow \phi_*(x)$;
- (c) if, moreover, ϕ is increasing, then so are $C_\varepsilon[\phi]$ and $C^\varepsilon[\phi]$.

Remark 2.4. If ϕ is increasing, then the sup or inf can be restricted to “one side” of x . More precisely, given $y \in \mathbb{R}^d$, define

$$\hat{y} = -(|y_1|, |y_2|, \dots, |y_d|),$$

Then $\hat{y} \leq y$ and $|\hat{y}| = |y|$, and so, because ϕ is increasing,

$$\phi(x - y) - \frac{|y|}{\varepsilon} \leq \phi(x - \hat{y}) - \frac{|\hat{y}|}{\varepsilon}.$$

It follows that

$$C^\varepsilon[\phi](x) = \sup_{y \geq 0} \left\{ \phi(x + y) - \frac{|y|}{\varepsilon} \right\},$$

and similarly

$$C_\varepsilon[\phi](x) = \inf_{y \leq 0} \left\{ \phi(x + y) + \frac{|y|}{\varepsilon} \right\}.$$

The second example involves convolving with mollifying functions that are weighted to one side. Let ρ be a smooth, positive function with support contained in $[-1, 1]^d$ and $\int \rho = 1$, and define

$$\rho_\varepsilon(z) = \frac{1}{\varepsilon^d} \rho\left(\frac{z - \varepsilon \mathbf{1}}{\varepsilon}\right) \quad \text{and} \quad \rho^\varepsilon(z) = \frac{1}{\varepsilon^d} \rho\left(\frac{z + \varepsilon \mathbf{1}}{\varepsilon}\right). \quad (2.4)$$

Lemma 2.5. *Assume $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$ is measurable and locally bounded, and set $\phi^\varepsilon := \rho^\varepsilon * \phi$ and $\phi_\varepsilon := \rho_\varepsilon * \phi$. Then, as $\varepsilon \rightarrow 0$, ϕ^ε and ϕ_ε converge to ϕ in L^p_{loc} for any $p \in [1, \infty)$. Moreover, if ϕ is increasing, then*

- (a) ϕ^ε and ϕ_ε are increasing, and $\phi_\varepsilon \leq \phi \leq \phi^\varepsilon$;
- (b) as $\varepsilon \rightarrow 0$, $\phi^\varepsilon \searrow \phi^*$ and $\phi_\varepsilon \nearrow \phi_*$;
- (c) for all $\varepsilon > 0$ and $x \in \mathbb{R}^d$, $\phi^\varepsilon(x - 2\varepsilon \mathbf{1}) = \phi_\varepsilon(x)$.

Proof. The convergence in L^p_{loc} is standard. For the rest of the proof, we assume ϕ is increasing. We see immediately that ϕ^ε and ϕ_ε are increasing. Observe that

$$\text{supp } \rho_\varepsilon \subset (0, 2\varepsilon)^d \quad \text{and} \quad \text{supp } \rho^\varepsilon \subset (-2\varepsilon, 0)^d.$$

As a consequence, because ϕ is increasing,

$$\phi^\varepsilon(x) = \int_{\mathbb{R}^d} \phi(x - y) \rho^\varepsilon(y) dy = \int_{(-2\varepsilon, 0)^d} \phi(x - y) \rho^\varepsilon(y) dy \geq \phi^*(x),$$

and similarly $\phi_\varepsilon(x) \leq \phi_*(x)$, so that part (a) is established.

Now assume that $0 < \varepsilon < \delta$. Using again that ϕ is increasing, we find

$$\begin{aligned} \phi^\delta(x) &= \frac{1}{\delta^d} \int_{(-2\delta, 0)^d} \phi(x - y) \rho\left(\frac{y}{\delta} + \mathbf{1}\right) dy \\ &\geq \frac{1}{\delta^d} \int_{(-2\delta, 0)^d} \phi\left(x - \frac{\varepsilon}{\delta} y\right) \rho\left(\frac{y}{\delta} + \mathbf{1}\right) dy \\ &= \frac{1}{\varepsilon^d} \int_{(-2\varepsilon, 0)^d} \phi(x - y) \rho\left(\frac{y}{\varepsilon} + \mathbf{1}\right) dy = \phi^\varepsilon(x), \end{aligned}$$

and similarly $\phi_\delta \leq \phi_\varepsilon$. The convergence statements in part (b) then easily follow.

Finally, part (c) is seen upon computing

$$\phi^\varepsilon(x - 2\varepsilon \mathbf{1}) = \frac{1}{\varepsilon^d} \int_{\mathbb{R}^d} \phi(y) \rho\left(\frac{x - y - \varepsilon \mathbf{1}}{\varepsilon}\right) dy = \phi_\varepsilon(x). \quad \blacksquare$$

Remark 2.5. An analogue of part (c) in Lemma 2.5 can also be seen for the sup- and inf-convolutions $C^\varepsilon[\phi]$ and $C_\varepsilon[\phi]$ in (2.3), namely, for all $R > 0$, there exists $C_R > 0$ depending on the linear growth of ϕ such that, for all $x \in B_R$,

$$C^\varepsilon[\phi](x - C_R \mathbf{1}) \leq C_\varepsilon[\phi](x).$$

The advantage of the one-sided mollifications $\phi * \rho^\varepsilon$ and $\phi * \rho_\varepsilon$ is that the constant C_R can be replaced by a uniform constant that does not depend on the growth of ϕ , which will be convenient when we consider ϕ depending on an additional time parameter in an L^1 way. On the other hand, the sup- and inf-convolutions are very flexible one-sided regularizations even when ϕ is not itself increasing.

2.4. A maximum principle

At various points in the paper, we apply a multi-dimensional maximum principle. For completeness, we provide a full proof, and we mention here the similarity in our methods with those used by Busca and Sirakov [15] and Capuzzo Dolcetta and Vitolo [16] to study nonlinear elliptic systems.

Lemma 2.6. *For some $M \in \mathbb{N}$, assume that*

$$\left\{ \begin{array}{l} (a^i)_{i=1}^M \subset C([0, T] \times \mathbb{R}^d, \mathbb{S}^d), (b^i)_{i=1}^M \subset C([0, T] \times \mathbb{R}^d, \mathbb{R}^d), \text{ and} \\ (c^i)_{i=1}^M, (d^i)_{i=1}^M, (\lambda_j^i)_{i,j=1}^M, (\mu_{k\ell}^i)_{i,k,\ell=1}^M \subset C([0, T] \times \mathbb{R}^d, \mathbb{R}) \\ \text{are uniformly bounded} \end{array} \right.$$

and

$$\left\{ \begin{array}{l} a^i(t, x) \geq 0, d^i(t, x) \geq 0, \lambda_j^i(t, x) \geq 0, \text{ and } \mu_{k\ell}^i(t, x) \leq 0 \\ \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d \text{ and } i, j, k, \ell \in \{1, 2, \dots, M\}. \end{array} \right.$$

Let $(V^i)_{i=1}^M \subset C^1([0, T] \times \mathbb{R}^d)$ be bounded on $[0, T] \times \mathbb{R}^d$ and satisfy the system

$$\begin{aligned} \partial_t V^i - \operatorname{tr}[a^i D^2 V^i] + b^i \cdot D V^i + c^i V^i + d^i \\ = \sum_{j \neq i} \lambda_j^i V^j + \sum_{k, \ell \neq i} \mu_{k\ell}^i V^k V^\ell \quad \text{in } (0, T) \times \mathbb{R}^d. \end{aligned} \quad (2.5)$$

Suppose that, for all $x \in \mathbb{R}^d$ and $i = 1, 2, \dots, M$, $V^i(0, x) \leq 0$. Then, for all $(t, x) \in (0, T] \times \mathbb{R}^d$ and $i = 1, 2, \dots, M$, $V^i(t, x) \leq 0$.

Proof. We prove the result first under the additional assumption that

$$\inf_{(t,x) \in [0,T] \times \mathbb{R}^d} d^i > 0, \quad \sup_{x \in \mathbb{R}^d} V^i(0, x) < 0,$$

and, for all $t \in [0, T]$ and $i \in \{1, 2, \dots, M\}$, $x \mapsto V^i(t, x)$ attains a global maximum on \mathbb{R}^d . In that case, define

$$t_0 := \inf \left\{ t \in (0, T] : \max_{x \in \mathbb{R}^d} \max_{i \in \{1, 2, \dots, M\}} V^i(t, x) = 0 \right\},$$

so that $t_0 > 0$. Let $x_0 \in \mathbb{R}^d$ and $i \in \{1, 2, \dots, M\}$ be such that the maximum is achieved. Note then that $D V^i(t_0, x_0) = 0$, $D^2 V^i(t_0, x_0) \leq 0$, $\partial_t V^i(t_0, x_0) \geq 0$, and, for all $j \in \{1, 2, \dots, M\}$ and $x \in \mathbb{R}^d$, $V^j(t_0, x) \leq V^i(t_0, x_0) \leq 0$. We thus obtain

$$d^i(t_0, x_0) \leq \sum_{j \neq i} \lambda_j^i V^j(t_0, x_0) + \sum_{k, \ell \neq i} \mu_{k\ell}^i V^k(t_0, x_0) V^\ell(t_0, x_0) \leq 0,$$

which is a contradiction in view of the assumption on d^i . In this case, we indeed conclude that $V^i(t, x) < 0$ for all $(t, x, i) \in [0, T] \times \mathbb{R}^d \times \{1, 2, \dots, M\}$.

We now turn to the general case. For $x \in \mathbb{R}^d$, define $v(x) := \sqrt{1 + |x|^2}$, and note that Dv and D^2v are globally bounded on \mathbb{R}^d . Then standard arguments yield that, for all $i \in \{1, 2, \dots, M\}$, $\beta > 0$, and $t \in [0, T]$, if S_β is the set of maximum points of

$$x \mapsto V^i(t, x) - \beta v(x),$$

then $\lim_{\beta \rightarrow 0} \beta \sup_{x \in S_\beta} v(x) = 0$. It follows that there exists a smooth bounded function $v_\beta: \mathbb{R}^d \rightarrow \mathbb{R}$ such that

$$\sup_{\beta > 0} (\|Dv_\beta\|_\infty + \|D^2v_\beta\|_\infty) < \infty, \quad \lim_{\beta \rightarrow 0} \beta \|v_\beta\|_\infty = 0,$$

and

$$\begin{aligned} & \max_{\substack{(t, x, i) \in [0, T] \times \mathbb{R}^d \\ \times \{1, 2, \dots, M\}}} (V^i(t, x) - \beta v(x)) \\ &= \max_{\substack{(t, x, i) \in [0, T] \times \mathbb{R}^d \\ \times \{1, 2, \dots, M\}}} (V^i(t, x) - \beta v_\beta(x)). \end{aligned}$$

Fix $\delta > 0$ and a constant $C > 0$ to be determined. For $(t, x) \in [0, T] \times \mathbb{R}^d$, $i \in \{1, 2, \dots, M\}$, set

$$\tilde{V}^i(t, x) := V^i(t, x) - \beta v_\beta(x) - \delta e^{Ct}.$$

Then, for all $i \in \{1, 2, \dots, M\}$, $\tilde{V}^i(0, \cdot) \leq -\delta$, and, for all $t \in [0, T]$, $\tilde{V}^i(t, \cdot)$ attains a global maximum over \mathbb{R}^d . Moreover, in $(0, T) \times \mathbb{R}^d$, $(\tilde{V}^i)_{i=1}^M$ solves the system

$$\partial_t \tilde{V}^i - \operatorname{tr}[a^i D^2 \tilde{V}^i] + b^i(t, x) \cdot D \tilde{V}^i(t, x) + c^i \tilde{V}^i + \tilde{d}^i = \sum_{j \neq i} \lambda_j^i \tilde{V}^j + \sum_{k, \ell \neq i} \mu_{k\ell}^i \tilde{V}^k \tilde{V}^\ell,$$

where

$$\begin{aligned} \tilde{d}^i(t, x) &= d^i(t, x) + C\delta e^{Ct} - \beta \operatorname{tr}[a^i D^2 v_\beta] + \beta b^i \cdot Dv_\beta \\ &\quad + \beta v_\beta(x) \left(c^i - \sum_{j \neq i} \lambda_j^i - \sum_{k, \ell \neq i} \mu_{k\ell}^i V^k \right) + \delta e^{Ct} \left(c^i - \sum_{j \neq i} \lambda_j^i - \sum_{k, \ell \neq i} \mu_{k\ell}^i V^k \right). \end{aligned}$$

From the boundedness of the coefficients and the V^i , and from the nonnegativity of d^i , it follows that there exists $\tilde{C} > 0$ depending only on the bounds of the coefficients and the V^i such that, as $\beta \rightarrow 0$,

$$\tilde{d}^i(t, x) \geq (C - \tilde{C})\delta e^{Ct} - o(1).$$

Taking $C > \tilde{C}$ and letting β be sufficiently small in relation to δ , we then see that $\tilde{d}^i > 0$. From the first step, we conclude that, if β is sufficiently small, then, for all $(t, x) \in [0, T] \times \mathbb{R}^d$ and $i \in \{1, 2, \dots, M\}$,

$$V^i(t, x) < \beta v_\beta(x) + \delta e^{Ct}.$$

Sending $\beta \rightarrow 0$ first and then $\delta \rightarrow 0$ yields the result. ■

3. The regular Lagrangian flow

The object of this section is to study the forward-in-time flow

$$\partial_t \phi_{t,s}(x) = b(t, \phi_{t,s}(x)) \quad \text{for } s \in [0, T], \quad t \in [s, T], \quad \phi_{s,s}(x) = x, \quad (3.1)$$

for a vector field $b: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfying

$$\begin{cases} \text{for some } C_0, C_1 \in L^1_+([0, T]), \\ |b(t, y)| \leq C_0(t)(1 + |y|) \text{ for a.e. } (t, y) \in [0, T] \times \mathbb{R}^d, \text{ and} \\ x \mapsto b(t, x) + C_1(t)x \text{ is increasing for a.e. } t \in [0, T]. \end{cases} \quad (3.2)$$

The second condition reads equivalently as $\partial_{x_j} b^i(t, \cdot) \geq -C_1(t)\delta_{ij}$ for all $i, j = 1, 2, \dots, d$. Moving forward, for convenience, we define the positive, increasing, absolutely continuous functions

$$\omega_0(t) = \int_0^t C_0(s) ds, \quad \omega_1(t) = \int_0^t C_1(s) ds, \quad t \in [0, T]. \quad (3.3)$$

Remark 3.1. Implicit in assumption (3.2) is a choice of basis on \mathbb{R}^d . Indeed, the results of the paper continue to hold if b is replaced by $Ab(t, A^\top x)$ for a $d \times d$ orthogonal matrix A .

The precise interpretation of problem (3.1), wherein b is discontinuous, is made sense of as a differential inclusion. Namely, at every discontinuity x of $b(t, \cdot)$, we have $b_*(t, x) \not\leq b^*(t, x)$, and so the natural formulation is for the absolutely continuous function $t \mapsto \phi_{t,s}(x)$ to satisfy

$$\partial_t \phi_{t,s}(x) \in [b_*(t, \cdot), b^*(t, \cdot)](\phi_{t,s}(x)) \quad \text{for } s \in [0, T], \text{ a.e. } t \geq s, \quad \phi_{s,s}(x) = x. \quad (3.4)$$

Remark 3.2. In order to make notation less cumbersome, for a function $\phi: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, we will always denote by ϕ_* and ϕ^* the lower- and upper-semicontinuous envelopes of ϕ in the space variable only; that is, for $(t, x) \in [0, T] \times \mathbb{R}^d$, $\phi_*(t, x) = \phi(t, \cdot)_*(x)$ and $\phi^*(t, x) = \phi(t, \cdot)^*(x)$.

If $\phi: \mathbb{R}^d \rightarrow \mathbb{R}^m$ for $m > 1$, then ϕ^* and ϕ_* are taken to be the coordinate-by-coordinate lower- and upper-semicontinuous envelopes, so, for instance, $\phi_* = (\phi_*^1, \phi_*^2, \dots, \phi_*^d)$; equivalently,

$$\phi_*(x) = \sup_{r>0} \inf_{|x-y|\leq r} \phi(y)$$

with respect to the partial order (2.1) on \mathbb{R}^m .

Remark 3.3. The differential inclusion (3.4), wherein $b(t, x)$ is replaced with the smallest box $[\alpha, \beta] = \prod_{i=1}^d [\alpha_i, \beta_i]$ containing all limit points of $b(t, y)$ as $y \rightarrow x$, is a slightly weaker formulation than the standard Filippov regularization [32], where boxes are replaced with general convex sets.

Before developing the general theory, it is useful to record the following a priori ODE bounds, which follow easily from Grönwall's lemma.

Lemma 3.1. *Assume, for some $C_0 \in L^1_+([0, T])$, that $b: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfies*

$$\operatorname{ess\,sup}_{x \in \mathbb{R}^d} \frac{|b(t, x)|}{1 + |x|} \leq C_0(t) \quad \text{for a.e. } t \in [0, T]$$

and X is an absolutely continuous solution of $\dot{X}(t) = b(t, X(t))$ in $[0, T]$ with $X(t_0) = x$ for some $t_0 \in [0, T]$. Then there exists a constant $C > 0$, depending only on $T > 0$, such that

$$|X(t) - X(s)| \leq C(1 + |x|) \int_s^t C_0(r) \, dr \quad \text{for } 0 \leq s \leq t \leq T.$$

3.1. A comparison principle

The following comparison result for ODEs is at the heart of much of the analysis in this paper. It leads to the existence and uniqueness of the regular Lagrangian flow, as well as stable notions of solutions to the transport and continuity equations with velocity fields satisfying (3.2).

Lemma 3.2. *Assume that $B: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfies*

$$\begin{cases} \{t \mapsto \|\nabla B(t, \cdot)\|_{L^\infty(B_R)}\} \in L^1([0, T]) \text{ for all } R > 0, \text{ and} \\ \partial_{x_i} B^j \geq 0 \text{ for all } i \neq j. \end{cases}$$

Let $X, Y \in C^{0,1}([0, T], \mathbb{R}^d)$ be such that, with respect to the partial order (2.1), $X(0) \leq Y(0)$ and

$$\dot{X}(t) \leq B(t, X(t)) \quad \text{and} \quad \dot{Y}(t) \geq B(t, Y(t)) \quad \text{for a.e. } t \in [0, T].$$

Then $X(t) \leq Y(t)$ for all $t \in [0, T]$.

Proof. The continuity of X and Y implies that there exists $R > 0$ such that $|X(t)| \vee |Y(t)| \leq R$ for all $t \in [0, T]$. We may thus assume without loss of generality that

$$\int_0^T \|\nabla B(t, \cdot)\|_\infty \, dt < \infty.$$

Define $\Delta = X - Y$ and, for $i, j = 1, 2, \dots, d$,

$$a_{ij}(t) = \int_0^1 \partial_{x_j} B^i(t, \tau X(t) + (1 - \tau)Y(t)) \, d\tau.$$

Observe that $\int_0^T |a_{ij}(t)| \, dt < \infty$ for all i, j , and $a_{ij} \geq 0$ whenever $i \neq j$. Then, for $i = 1, 2, \dots, d$,

$$\Delta^i(0) \leq 0 \quad \text{and} \quad \dot{\Delta}^i(t) \leq a_{ii}(t)\Delta^i(t) + \sum_{i \neq j} a_{ij}(t)\Delta^j(t) \quad \text{for } t \in [0, T].$$

Fix $\delta > 0$, set

$$\psi(t) = t + \int_0^t \max_{i=1,2,\dots,d} \sum_{j=1}^d |a_{ij}(s)| ds,$$

and define

$$t_0 := \inf\{t \in (0, T] : \text{there exists } i \in \{1, 2, \dots, d\} \text{ such that } e^{-\psi(t)} \Delta^i(t) > \delta\}.$$

Since $e^{-\psi(0)} \Delta^i(0) \leq 0 < \delta$, we must have $t_0 > 0$. Assume by contradiction that $t_0 < \infty$. Then, by continuity, there exists i such that, for all j and $s \in [0, t_0]$,

$$e^{-\psi(s)} \Delta^j(s) \leq e^{-\psi(t_0)} \Delta^i(t_0) = \delta.$$

We compute

$$\begin{aligned} 0 &\leq \frac{d}{dt} e^{-\psi(t)} \Delta^i(t) \Big|_{t=t_0} \\ &\leq -\dot{\psi}(t_0) e^{-\psi(t_0)} \Delta^i(t_0) + e^{-\psi(t_0)} \left(a_{ii}(t_0) \Delta^i(t_0) + \sum_{j \neq i} a_{ij}(t_0) \Delta^j(t_0) \right) \\ &\leq -\dot{\psi}(t_0) e^{-\psi(t_0)} \Delta^i(t_0) + e^{-\psi(t_0)} \left(a_{ii}(t_0) \Delta^i(t_0) + \sum_{j \neq i} a_{ij}(t_0) \Delta^i(t_0) \right) \\ &= -\delta \left(\dot{\psi}'(t_0) - \sum_{j=1}^d a_{ij}(t_0) \right) \\ &\leq -\delta, \end{aligned}$$

which is a contradiction. It follows that $\Delta^i(t) < \delta e^{\psi(t)}$ for all $t \in [0, T]$ and $i = 1, 2, \dots, d$. Sending $\delta \rightarrow 0$ yields the result. ■

3.2. Maximal and minimal flows

The comparison principle from the previous subsection is now used to establish the existence of maximal and minimal (with respect to the order (2.1)) semicontinuous solutions of the differential inclusion (3.4) for vector fields satisfying $\partial_{x_i} b^j \geq 0$ for $i \neq j$.

Proposition 3.1. *Assume that $b: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfies*

$$\int_0^T \operatorname{ess\,sup}_{y \in \mathbb{R}^d} \frac{|b(t, y)|}{1 + |y|} dt < \infty \quad \text{and} \quad \partial_{x_i} b^j \geq 0 \quad \text{for all } i \neq j.$$

Then there exist solutions ϕ^+ and ϕ^- of (3.4) that are absolutely continuous in time such that

- (a) *for all $0 \leq s \leq t \leq T$, $\phi_{t,s}^+$ and $\phi_{t,s}^-$ are increasing in the sense of (2.1), $\phi_{t,s}^+$ is right-continuous, and $\phi_{t,s}^-$ is left-continuous;*

- (b) if ϕ is any other solution of (3.4), then $\phi^- \leq \phi \leq \phi^+$;
- (c) if $0 \leq r \leq s \leq t \leq T$, then $\phi_{t,s}^+ \circ \phi_{s,r}^+ = \phi_{t,r}^+$ and $\phi_{t,s}^- \circ \phi_{s,r}^- = \phi_{t,r}^-$;
- (d) for all $0 \leq s \leq t \leq T$, $(\phi_{t,s}^+)^* \geq \phi_{t,s}^-$.

Remark 3.4. When $d = 1$, the condition on the derivatives of b becomes vacuous, and we recover the existence of unique minimal and maximal solutions under the sole assumption that b is locally bounded.

Proof of Proposition 3.1. Define $M \in L_+^1([0, T])$ by $M = \text{ess sup}_{y \in \mathbb{R}^d} \frac{|b(\cdot, y)|}{1+|y|}$, and, for $\varepsilon \in (0, 1)$, define b^ε and b_ε component by component by

$$b^{i,\varepsilon}(t, \cdot) = C^{\varepsilon M(t)^{-1}}[b^i(t, \cdot)] \quad \text{and} \quad b_\varepsilon^i(t, \cdot) = C_{\varepsilon M(t)^{-1}}[b^i(t, \cdot)],$$

where C_ε and C^ε are the sup- and inf-convolution operators as in (2.3). By Lemma 2.4, b^ε and b_ε are well defined, and, for a.e. t , $|\nabla b^\varepsilon(t, \cdot)| \vee |\nabla b_\varepsilon(t, \cdot)| \leq \varepsilon^{-1} M(t)$, so that, for all $(s, x) \in [0, T] \times \mathbb{R}^d$, there exist unique solutions $\phi^{\pm, \varepsilon}$ of

$$\partial_t \phi_{t,s}^{+, \varepsilon}(x) = b^\varepsilon(t, \phi_{t,s}^{+, \varepsilon}(x))$$

and

$$\partial_t \phi_{t,s}^{-, \varepsilon}(x) = b_\varepsilon(t, \phi_{t,s}^{-, \varepsilon}(x))$$

in $[s, T]$, with $\phi_{s,s}^{+, \varepsilon}(x) = \phi_{s,s}^{-, \varepsilon}(x) = x$. By Lemma 3.1, for every $(s, x) \in [0, T] \times \mathbb{R}^d$, $\phi_{s,s}^{+, \varepsilon}(x)$ and $\phi_{s,s}^{-, \varepsilon}(x)$ are bounded and continuous uniformly in ε .

The comparison result Lemma 3.2 immediately gives $\phi^{-, \varepsilon} \leq \phi^{+, \varepsilon}$, and, if $x \leq y$, $\phi_{t,s}^{-, \varepsilon}(x) \leq \phi_{t,s}^{-, \varepsilon}(y)$ and $\phi_{t,s}^{+, \varepsilon}(x) \leq \phi_{t,s}^{+, \varepsilon}(y)$.

Fix $0 < \varepsilon < \delta$. Then Lemma 2.4(b) implies that $b^\varepsilon \leq b^\delta$ and $b_\varepsilon \geq b_\delta$. It follows that

$$\partial_t \phi_{t,s}^{+, \varepsilon} \leq b^\delta(t, \phi_{t,s}^{+, \varepsilon}),$$

and so Lemma 3.2 yields $\phi^{+, \varepsilon} \leq \phi^{+, \delta}$. A similar argument gives $\phi^{-, \varepsilon} \geq \phi^{-, \delta}$. It now follows that there exist bounded and Lipschitz functions $\phi_{\cdot, s}^+(x)$ and $\phi_{\cdot, s}^-(x)$ on $[s, T]$, which are increasing in x , such that, as $\varepsilon \rightarrow 0$,

$$\phi_{t,s}^{+, \varepsilon}(x) \searrow \phi_{t,s}^+(x) \quad \text{and} \quad \phi_{t,s}^{-, \varepsilon}(x) \nearrow \phi_{t,s}^-(x) \quad \text{uniformly for } t \in [s, T].$$

Moreover, for $0 \leq s \leq t_1 \leq t_2 \leq T$, $x \in \mathbb{R}^d$, and $0 < \varepsilon < \delta$, we have

$$\phi_{t_2, s}^{+, \varepsilon}(x) \leq \phi_{t_2, s}^{+, \varepsilon}(x) + \int_{t_1}^{t_2} b^\delta(r, \phi_{r, s}^{+, \varepsilon}(x)) dr,$$

and so, sending $\varepsilon \rightarrow 0$ and using that $\nabla b^\delta \in L_+^1([0, T], L^\infty(\mathbb{R}^d))$, we find

$$\phi_{t_2, s}^+(x) \leq \phi_{t_2, s}^+(x) + \int_{t_1}^{t_2} b^\delta(r, \phi_{r, s}^+(x)) dr.$$

We may replace b^δ by b^* upon sending $\delta \rightarrow 0$ and appealing to the monotone convergence theorem, and so, because t_1 and t_2 were arbitrary, this yields $\partial_t \phi_{t,s}^+(x) \leq b^*(t, \phi_{t,s}^+(x))$. Arguing similarly establishes that $\partial_t \phi_{t,s}^+(x) \geq b_*(t, \phi_{t,s}^+(x))$, and we conclude that ϕ^+ is a solution of (3.4). An analogous argument proves that ϕ^- solves (3.4) as well.

Fix $0 \leq s \leq t \leq T$ and $x \in \mathbb{R}^d$ again. If $h \geq 0$, then, for arbitrary $\varepsilon > 0$,

$$\phi_{t,s}^+(x) \leq \phi_{t,s}^+(x+h) \leq \phi_{t,s}^{+,\varepsilon}(x+h).$$

Sending $h \searrow 0$ first and then $\varepsilon \rightarrow 0$ gives $\lim_{h \searrow 0} \phi_{t,s}^+(x+h) = \phi_{t,s}^+(x)$, so that $\phi_{t,s}^+$ is right-continuous. A similar argument shows that $\phi_{t,s}^-(x)$ is left-continuous, and this completes the proof of part (a).

We next prove part (b). If ϕ is any other solution of (3.4), then, for all $\varepsilon > 0$,

$$b_\varepsilon(t, \phi_{t,s}(x)) \leq \partial_t \phi_{t,s}(x) \leq b^\varepsilon(t, \phi_{t,s}(x)) \quad \text{for } t \in [s, T] \text{ and } \phi_{s,s}(x) = x.$$

Appealing once more to Lemma 3.2 gives $\phi_{t,s}^{-,\varepsilon}(x) \leq \phi_{t,s}(x) \leq \phi_{t,s}^{+,\varepsilon}(x)$, and sending $\varepsilon \rightarrow 0$ gives the result.

We now prove the flow property stated in part (c). Suppose $x \in \mathbb{R}^d$ and $0 \leq r \leq s \leq t \leq T$. Then, for $\varepsilon > 0$, because $\phi_{t,s}^{+,\varepsilon}$ is increasing,

$$\phi_{t,s}^{+,\varepsilon}(\phi_{s,r}^+(x)) \leq \phi_{t,s}^{+,\varepsilon}(\phi_{s,r}^{+,\varepsilon}(x)) = \phi_{t,r}^{+,\varepsilon}(x).$$

Taking $\varepsilon \rightarrow 0$ yields $\phi_{t,s}^+(\phi_{s,r}^+(x)) \leq \phi_{t,r}^+(x)$. For the opposite inequality, observe that $[s, T] \ni t \mapsto \phi_{t,r}^+(x)$ solves (3.4) with value $\phi_{s,r}^+(x)$ at time $t = s$. It follows from part (b) that $\phi_{t,r}^+(x) \leq \phi_{t,s}^+(\phi_{s,r}^+(x))$. The argument for the flow property for ϕ^- is analogous.

Let $x \in \mathbb{R}^d$. Sending $y \nearrow x$ in the inequality $\phi_{t,s}^-(y) \leq \phi_{t,s}^+(y)$, using the fact that ϕ^- and ϕ^+ are increasing and ϕ^- is left-continuous, yields $\phi_{t,s}^-(x) \leq (\phi_{t,s}^+)_*(x)$. This concludes the proof of part (d). ■

3.3. Uniqueness and stability of the regular Lagrangian flow

We now use the full assumption (3.2), and in particular, the lower bounds on $\partial_{x_i} b^i$ for $i = 1, 2, \dots, d$ that were not needed to prove the existence of the maximal and minimal solution of (3.4). In particular, we prove that ϕ^+ and ϕ^- are equal almost everywhere, giving rise to a unique regular Lagrangian flow.

Proposition 3.2. *Assume that b satisfies (3.2). If ϕ^+ and ϕ^- are the maximal and minimal flow from Proposition 3.1, then, for all $0 \leq s \leq t \leq T$, $(\phi_{t,s}^+)_* = \phi_{t,s}^-$. Moreover, for all $(s, x) \in [0, T] \times \mathbb{R}^d$ and a.e. $t \in [s, T]$,*

$$\partial_t \phi_{t,s}^+(x) = b^*(t, \phi_{t,s}^+(x)) \quad \text{and} \quad \partial_t \phi_{t,s}^-(x) = b_*(t, \phi_{t,s}^-(x)).$$

Proof. Let

$$\tilde{\phi}_{t,s}^\pm(x) = e^{\omega_1(t)} \phi_{t,s}^\pm(e^{-\omega_1(t)} x)$$

and

$$\tilde{b}(t, x) = e^{\omega_1(t)} b(t, e^{-\omega_1(t)} x) + C_1(t)x,$$

where ω_1 is as in (3.3). Then \tilde{b} satisfies (3.2) with $C_1 = 0$ and with a possibly different $C_0 \in L^1_+$, and $\tilde{\phi}^\pm$ are the corresponding maximal and minimal flows. We may therefore assume without loss of generality that $b(t, \cdot)$ is increasing for $t \in [0, T]$.

We first prove the statement involving the lower-semicontinuous envelope of $\phi_{t,s}^+$, and in view of Proposition 3.1 (d), we need only prove the opposite inequality $(\phi_{t,s}^+)_* \leq \phi_{t,s}^-$. For $t \in [0, T]$, define

$$b^\varepsilon(t, \cdot) = b(t, \cdot) * \rho^\varepsilon \quad \text{and} \quad b_\varepsilon(t, \cdot) = b(t, \cdot) * \rho_\varepsilon,$$

where the one-sided mollifiers ρ^ε and ρ_ε are defined in (2.4), and let $\phi^{+,\varepsilon}$ and $\phi^{-,\varepsilon}$ be the corresponding flows. Arguing exactly as in the proof of Proposition 3.1, appealing to Lemma 2.5, as $\varepsilon \rightarrow 0$, $\phi_{t,s}^{+,\varepsilon}(x) \searrow \phi_{t,s}^+(x)$ and $\phi_{t,s}^{-,\varepsilon}(x) \nearrow \phi_{t,s}^-(x)$ uniformly on $[s, T]$, for any $s \in [0, T]$ and $x \in \mathbb{R}^d$.

Fix $0 < \varepsilon < \delta$ and define $X^{\delta,\varepsilon}(t) = \phi_{t,s}^{+,\varepsilon}(x - 2\delta\mathbf{1}) + 2\delta\mathbf{1}$. Then, in view of Lemma 2.5 (c), $X^{\delta,\varepsilon}$ satisfies

$$\dot{X}^{\delta,\varepsilon}(t) = b^\varepsilon(t, X^{\delta,\varepsilon}(t) - 2\delta\mathbf{1}) \leq b_\varepsilon(t, X^{\delta,\varepsilon}(t)) \quad \text{in } [s, T] \quad \text{and} \quad X^{\delta,\varepsilon}(s) = x.$$

We may thus appeal to Lemma 3.2, and find that, for all $t \in [s, T]$,

$$\phi_{t,s}^{+,\varepsilon}(x - 2\delta\mathbf{1}) \leq \phi_{t,s}^{-,\varepsilon}(x) - 2\delta\mathbf{1}.$$

Sending $\varepsilon \rightarrow 0$ first and then $\delta \rightarrow 0$ gives $(\phi_{t,s}^+)_* \leq \phi_{t,s}^-$, as desired.

We now prove the final statement. For $0 < \varepsilon < \delta$, we have the inequalities, for $t \in [s, T]$ and $x \in \mathbb{R}^d$,

$$b^*(t, \phi_{t,s}^+(x)) \leq b^*(t, \phi_{t,s}^{+,\varepsilon}(x)) \leq b^\varepsilon(t, \phi_{t,s}^{+,\varepsilon}(x)) \leq b^\delta(t, \phi_{t,s}^{+,\varepsilon}(x)).$$

Since b^δ is Lipschitz continuous in the space variable,

$$\lim_{\varepsilon \rightarrow 0} b^\delta(t, \phi_{t,s}^{+,\varepsilon}(x)) = b^\delta(t, \phi_{t,s}^+(x)).$$

Therefore,

$$b^*(t, \phi_{t,s}^+(x)) \leq \liminf_{\varepsilon \rightarrow 0} b^\varepsilon(t, \phi_{t,s}^{+,\varepsilon}(x)) \leq \limsup_{\varepsilon \rightarrow 0} b^\varepsilon(t, \phi_{t,s}^{+,\varepsilon}(x)) \leq b^\delta(t, \phi_{t,s}^+(x)),$$

and so, sending $\delta \rightarrow 0$, it follows that

$$\lim_{\varepsilon \rightarrow 0} b^\varepsilon(t, \phi_{t,s}^{+,\varepsilon}(x)) = b^*(t, \phi_{t,s}^+(x)).$$

For arbitrary $\varepsilon > 0$, $0 \leq s \leq t \leq T$, and $x \in \mathbb{R}^d$,

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

Sending $\varepsilon \rightarrow 0$ and appealing to dominated convergence gives

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

and a similar argument gives

$$\phi_{t,s}^-(x) = x + \int_s^t b_*(r, \phi_{r,s}^-(x)) dr. \quad \blacksquare$$

Recalling that increasing functions are continuous almost everywhere, it now follows from Proposition 3.2 that $\phi_{t,s}^+$ and $\phi_{t,s}^-$ are equal almost everywhere, given $0 \leq s \leq t \leq T$. We thus finally arrive at the almost-everywhere unique solvability of ODE (3.1), and the identification of the unique regular Lagrangian flow.

Theorem 3.1. *For every $s \in [0, T]$ and for almost every $x \in \mathbb{R}^d$, there exists a unique absolutely continuous solution $[s, T] \ni t \mapsto \phi_{t,s}(x)$ of the differential inclusion (3.4). Moreover, $\phi_{t,s}$ satisfies the regular Lagrangian property: the map $C_c(\mathbb{R}^d) \ni f \mapsto f \circ \phi_{t,s}$ extends continuously to $f \in L_{\text{loc}}^1(\mathbb{R}^d)$, and there exists $C > 0$ depending only on T such that, for all $R > 0$, $f \in L_{\text{loc}}^1$, $t \in [s, T]$,*

$$\|f \circ \phi_{t,s}\|_{L^1(B_R)} \leq e^{d(\omega_1(t) - \omega_1(s))} \|f\|_{L^1(B_{R+C(1+R)(\omega_0(t) - \omega_0(s))})}. \quad (3.5)$$

In particular,

$$\{[s, T] \times \mathbb{R}^d \ni (t, x) \mapsto b(t, \phi_{t,s}(x)) \in L_{\text{loc}}^1(\mathbb{R}^d, L^1([s, T]))\},$$

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

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

Finally, if $0 \leq r \leq s \leq t \leq T$, then the composition $\phi_{t,s} \circ \phi_{s,r}$ is well defined a.e. and is equal to $\phi_{t,r}$.

Remark 3.5. Taking $f = \mathbf{1}_A$ in the estimate above, for some $A \subset \mathbb{R}^d$ of finite measure, we find the regular Lagrange property

$$|\phi_{t,s}^{-1}(A)| \leq e^{d(\omega_1(t) - \omega_1(s))} |A|.$$

Proof of Theorem 3.1. Let $s \in [0, T]$ be fixed and, for $N \in \mathbb{N}$, define $t_n^N = s + \frac{n(T-s)}{N}$, $n = 0, 1, 2, \dots, N$. Then, by Lemma 2.3 and Proposition 3.2, there exist sets $B_n^N \subset \mathbb{R}^d$ of full measure such that $\phi_{t_n^N, s}^+ = \phi_{t_n^N, s}^-$ on B_n^N . Now define the full measure set

$$B = \bigcap_{N \in \mathbb{N}} \bigcap_{n=0}^N B_n^N.$$

Let $(t, x) \in [s, T] \times B$, and, for any $N \in \mathbb{N}$, let $n(t) = 0, 1, 2, \dots, N$ be such that $|t_n^N - t|$ is minimized. Then, in view of the uniform continuity in time of ϕ^+ and ϕ^- , for some $\omega_N > 0$ with $\lim_{N \rightarrow \infty} \omega_N = 0$,

$$|\phi_{t,s}^+(x) - \phi_{t,s}^-(x)| \leq |\phi_{t_n(t),s}^+(x) - \phi_{t,s}^+(x)| + |\phi_{t_n(t),s}^-(x) - \phi_{t,s}^-(x)| \leq \omega_N,$$

and, since N was arbitrary, we find that $\phi_{t,s}^+(x) = \phi_{t,s}^-(x)$. It follows that, up to a full measure set, $\phi_{\cdot,s}^+$ and $\phi_{\cdot,s}^-$ may be identified as absolutely continuous functions on $[s, T]$.

For $f \in L^1(\mathbb{R}^d) \cap C_c(\mathbb{R}^d)$, let $\varepsilon > 0$ and let b^ε and $\phi^{+, \varepsilon}$ be as in the proof of Proposition 3.1. Then

$$\partial_t \det(D_x \phi_{t,s}^{+, \varepsilon}(x)) = \operatorname{div} b^\varepsilon(t, \phi_{t,s}^{+, \varepsilon}(x)) \det(D_x \phi_{t,s}^{+, \varepsilon}(x)) \geq -d C_1(t) \det(D_x \phi_{t,s}^{+, \varepsilon}(x)),$$

from which it follows that $\det(D_x \phi_{t,s}^{+, \varepsilon}(x)) \geq \exp(-d \int_s^t C_1(r) dr)$. The change of variables formula then gives

$$\int_{\mathbb{R}^d} |f(\phi_{t,s}^{+, \varepsilon}(x))| dx \leq \exp\left(d \int_s^t C_1(r) dr\right) \int_{\mathbb{R}^d} |f(x)| dx.$$

The set $(\phi_{t,s}^{+, \varepsilon}(x))_{x \in \operatorname{supp} f, \varepsilon > 0}$ is uniformly bounded in view of Lemma 3.1, and so the bounded convergence theorem implies, upon taking $\varepsilon \rightarrow 0$, that

$$\int_{\mathbb{R}^d} |f(\phi_{t,s}(x))| dx \leq \exp\left(d \int_s^t C_1(r) dr\right) \int_{\mathbb{R}^d} |f(x)| dx.$$

This implies that the map $f \mapsto f \circ \phi_{t,s}$ extends continuously to $L^1(\mathbb{R}^d)$. The local statement follows from the finite speed of propagation implied by the a priori estimates in Lemma 3.1.

It now follows easily that $b(\cdot, \phi_{\cdot,s})$ belongs to $L_{\operatorname{loc}}^1([0, T] \times \mathbb{R}^d)$, and the uniqueness statement for absolutely continuous solutions of the integral equation follows from Proposition 3.2 and the fact that $\phi = \phi^+ = \phi^-$ a.e.

Finally, the composition $\phi_{t,s} \circ \phi_{s,r}$ is justified because $\phi_{t,s} \in L_{\operatorname{loc}}^1$, and its equality to $\phi_{t,r}$ a.e. is a consequence of the a.e. uniqueness of the ODE and the flow properties in Proposition 3.1 (c). ■

We now demonstrate that any regularizations of b lead to the flow ϕ in the limit, not only the one-sided regularizations used above.

Theorem 3.2. Assume $(b^\varepsilon)_{\varepsilon > 0}$ is a family satisfying (3.2) uniformly in ε ; for every ε ,

$$\int_0^T \|\nabla b^\varepsilon(t, \cdot)\|_\infty dt < \infty,$$

and, as $\varepsilon \rightarrow 0$, b^ε converges a.e. to b . If ϕ^ε is the unique flow corresponding to b^ε , then, for all $s \in [0, T]$ and $p \in [1, \infty)$, as $\varepsilon \rightarrow 0$, $\phi^\varepsilon \rightarrow \phi$ in $C([s, T], L_{\operatorname{loc}}^p(\mathbb{R}^d))$ and in $L_{\operatorname{loc}}^p(\mathbb{R}^d, W^{1,1}([s, T]))$.

Proof. The a priori bounds in Lemma 3.1 imply that $(\phi_{\cdot,s}^\varepsilon)_{\varepsilon>0}$ is bounded uniformly in $L_{\text{loc}}^\infty([s, T] \times \mathbb{R}^d)$, and in particular, for all $p \in [1, \infty)$ and $R > 0$,

$$\sup_{\varepsilon>0} \sup_{t \in [s, T]} \int_{B_R} |\phi_{t,s}^\varepsilon(x)|^p dx < \infty.$$

Moreover, for all $t \in [s, T]$, $\phi_{t,s}^\varepsilon$ is increasing, and so, for all $R > 0$, there exist C and C' depending on d and R such that

$$\int_{B_R} |\nabla \phi_{t,s}^\varepsilon(x)| dx \leq \sum_{i=1}^d \int_{B_R} \partial_{x_i} \phi_{t,s}^\varepsilon(x) dx \leq C \|\phi_{t,s}^\varepsilon\|_{L^\infty(B_R)} \leq C'.$$

Using the uniform L_{loc}^∞ bound once more, this implies that, for all $0 \leq h \leq 1$, $t \in [s, T]$, and some $C'' > 0$ depending on R and p ,

$$\int_{B_R} |\phi_{t,s}^\varepsilon(x+h) - \phi_{t,s}^\varepsilon(x)|^p dx \leq C''h,$$

from which we conclude that there exists a common compact subset $K \subset L_{\text{loc}}^p$ such that $\phi_{t,s}^\varepsilon \in K$ for all $\varepsilon > 0$ and $t \in [s, T]$. Finally, the a priori estimates in Lemma 3.1 also give (for C as in the statement of that lemma)

$$\int_{B_R} |\phi_{t_1,s}^\varepsilon(x) - \phi_{t_2,s}^\varepsilon(x)|^p dx \leq C^p(1+R)^p |B_R| \int_{t_1}^{t_2} C_0(r) dr.$$

The Arzelà–Ascoli theorem thus implies that $(\phi_{\cdot,s}^\varepsilon)_{\varepsilon>0}$ is precompact in $C([s, T], L_{\text{loc}}^p(\mathbb{R}^d))$, and therefore, for some subsequence $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$,

$$\phi_{\cdot,s}^{\varepsilon_n} \text{ converges as } n \rightarrow \infty \text{ in } C([s, T], L_{\text{loc}}^p(\mathbb{R}^d)) \text{ to some } \psi_{\cdot,s},$$

and then $\psi_{\cdot,s}$ also satisfies the regular Lagrangian property (3.5).

We now use (3.5) to deduce that, for any $R > 0$, there exists $R' > 0$ independent of ε such that, for a.e. $r \in [s, T]$ and for any $\delta > \varepsilon$,

$$\begin{aligned} & \|b^\varepsilon(r, \phi_{r,s}^\varepsilon) - b(r, \psi_{r,s})\|_{L^p(B_R)} \\ & \leq \|b^\varepsilon(r, \phi_{r,s}^\varepsilon) - b^\delta(r, \phi_{r,s}^\varepsilon)\|_{L^p(B_R)} + \|b^\delta(r, \phi_{r,s}^\varepsilon) - b^\delta(r, \psi_{r,s})\|_{L^p(B_R)} \\ & \quad + \|b^\delta(r, \psi_{r,s}) - b(r, \psi_{r,s})\|_{L^p(B_R)} \\ & \leq e^{d(\omega_1(t) - \omega_1(s))} (\|b^\varepsilon(r, \cdot) - b^\delta(r, \cdot)\|_{L^p(B_{R'})} + \|b^\delta(r, \cdot) - b(r, \cdot)\|_{L^p(B_{R'})}) \\ & \quad + \|b^\delta(r, \phi_{r,s}^\varepsilon) - b^\delta(r, \psi_{r,s})\|_{L^p(B_R)}. \end{aligned}$$

Taking $\varepsilon_n \rightarrow 0$ first (using the Lipschitz continuity of b^δ) and then $\delta \rightarrow 0$, we find that the right-hand side above converges to zero. By Lemma 3.1, we may use the dominated convergence theorem to deduce that

$$\lim_{n \rightarrow \infty} \int_s^T \|b^{\varepsilon_n}(r, \phi_{r,s}^{\varepsilon_n}) - b(r, \psi_{r,s})\|_{L^p(B_R)} dr = 0.$$

Sending $n \rightarrow \infty$ in the integral equation

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

thus gives, in the distributional sense,

$$\psi_{t,s}(x) = x + \int_s^t b(r, \psi_{r,s}(x)) dr. \quad (3.6)$$

We also have, by Minkowski's inequality,

$$\|\partial_t \phi_{s,\cdot}^{\varepsilon_n} - \partial_t \psi_{s,\cdot}\|_{L^p(B_R, L^1([s, T]))} \leq \int_s^T \|b^{\varepsilon_n}(r, \phi_{r,s}^{\varepsilon_n}) - b(r, \psi_{r,s})\|_{L^p(B_R)} dr \xrightarrow{n \rightarrow \infty} 0,$$

and thus (3.6) is satisfied in the integral sense for a.e. $x \in \mathbb{R}^d$. By Theorem 3.1, $\psi = \phi$, and therefore the convergence statements hold for the full family $\varepsilon \rightarrow 0$. ■

Another means of regularization is through the addition of stochastic noise. In particular, if $s \in [0, T]$, $x \in \mathbb{R}^d$, $\varepsilon > 0$, and $W: \Omega \times [0, T] \rightarrow \mathbb{R}^d$ is a Wiener process defined over a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then there exists a unique strong (that is, adapted to the natural filtration of W) solution of the SDE

$$d_t \phi_{t,s}^\varepsilon(x) = b(t, \phi_{t,s}^\varepsilon(x)) dt + \sqrt{2\varepsilon} dW_t. \quad (3.7)$$

This is true even if b is merely locally bounded; see [25, 59].

The following is then proved exactly as for Theorem 3.2.

Theorem 3.3. *Assume b satisfies (3.2) and let ϕ^ε be the solution of (3.7). Then, for all $s \in [0, T]$, with probability 1, as $\varepsilon \rightarrow 0$, $\phi_{\cdot,s}^\varepsilon$ converges in $C([s, T], L_{\text{loc}}^p(\mathbb{R}^d))$ and $L_{\text{loc}}^p(\mathbb{R}^d, C([s, T]))$ to $\phi_{\cdot,s}$.*

Proof. The transformation $\tilde{\phi}_{t,s}(x) = \phi_{t,s}(x) - \sqrt{2\varepsilon}(W_t - W_s)$ leads to the equation

$$\partial_t \tilde{\phi}_{t,s}^\varepsilon(x) = b(t, \tilde{\phi}_{t,s}^\varepsilon(x) + \sqrt{2\varepsilon}(W_t - W_s)).$$

Exactly the same arguments as those above show that $\tilde{\phi}_{t,s}^\varepsilon$ is increasing for all $0 \leq s \leq t \leq T$, satisfies the regular Lagrange property (3.5), and, for any fixed Brownian path W , the a priori estimates of Lemma 3.1 may be applied. Arguing as in the proof of Theorem 3.2, we may then extract a subsequence $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$ such that $\tilde{\phi}^{\varepsilon_n}$, and therefore ϕ^{ε_n} , converges in $C([s, T], L_{\text{loc}}^p(\mathbb{R}^d))$ and $L_{\text{loc}}^p(\mathbb{R}^d, C([s, T]))$. The same arguments as in Theorem 3.2 may then be used to conclude that any such limit is the unique regular Lagrangian flow ϕ from Theorem 3.1. ■

If b is smooth, then $\partial_s \phi_{t,s}(x) = -b(s, \phi_{t,s}(x))$. It is then straightforward to show that all the same theory as above can be developed for the terminal value problem for the flow corresponding to $-b$.

Theorem 3.4. *For every $t \in [0, T]$ and $p \in [1, \infty)$, $\phi_{t,\cdot} \in C([0, t], L^p_{\text{loc}}(\mathbb{R}^d)) \cap L^p_{\text{loc}}(\mathbb{R}^d, W^{1,1}([0, t]))$, and, for almost every $x \in \mathbb{R}^d$, $[0, t] \ni s \mapsto \phi_{t,s}(x)$ is the unique absolutely continuous solution of*

$$\phi_{t,s}(x) = x + \int_s^t b(r, \phi_{t,r}(x)) dr, \quad s \in [0, t].$$

3.4. Some remarks on the backward flow

We discuss next the question of solvability of the backward flow: for fixed $s \in [0, T]$, this is the terminal value problem

$$\partial_t \phi_{t,s}(x) = b(t, \phi_{t,s}(x)), \quad t \in [0, s], \quad \phi_{s,s}(x) = x. \quad (3.8)$$

Formally, the lower bound on $\text{div } b$ suggests that the backward flow should concentrate positive measure sets to null sets. As an example when $d = 1$, if $b(t, x) = \text{sgn } x$, then the backward flow is given by

$$\phi_{t,s}(x) = x - (\text{sgn } x)(s - t) \quad \text{for } x \in \mathbb{R}, \quad t < s,$$

and so all trajectories eventually concentrate at $x = 0$. This situation can be generalized to multiple dimensions when b satisfies the half-Lipschitz property

$$(b(t, x) - b(t, y)) \cdot (x - y) \geq -C_1(t)|x - y|^2, \quad C_1 \in L^1_+[0, T]. \quad (3.9)$$

Condition (3.9) then implies the existence of a unique, Lipschitz, concentrating solution of the backward flow (3.8). This condition was used by Filippov [32] to build unique solutions of differential inclusions; see also [23, 53, 54] and the recent work of the authors [48].

The situation is very different when b satisfies (3.2). The differential inclusion corresponding to (3.8) then takes the form

$$-\partial_t \phi_{t,s}(x) \in [-b^*(t, \cdot), -b_*(t, \cdot)](\phi_{t,s}(x)), \quad s \in [0, T], \quad 0 \leq t \leq s, \quad \phi_{s,s}(x) = x. \quad (3.10)$$

As it turns out, there is no unique flow, and, in some cases, no Lipschitz flow.

Example 3.1. For $x = (x_1, x_2) \in \mathbb{R}^2$, consider the vector field

$$b(x) = \text{sgn } x_1(1, 1).$$

Since b is independent of t , we formally write the flow $\phi_{t,s}$ for $0 \leq t \leq s \leq T$ as ϕ_{t-s} .

Observe that, for all $x_2 \in \mathbb{R}$, $b_*(0, x_2) = -(1, 1)$ and $b^*(0, x_2) = (1, 1)$. It then follows that the direction in which $\phi_{-\tau}(0, x_2)$ moves for $\tau > 0$ lies in the box $[-1, 1]^2$, and so any flow that solves (3.10) must take the form, for some $c \in [-1, 1]$ and all $\tau \geq 0$,

$$\phi_{-\tau}^{(c)}(x) = \begin{cases} x - (\text{sgn } x_1)(1, 1)\tau, & |x_1| \geq \tau, \\ (0, x_2 + c(\tau - |x_1|)), & |x_1| < \tau. \end{cases}$$

There is therefore no unique flow in the strip $\{|x_1| < \tau\}$, even under the restriction that the flow be Lipschitz.

Recall that (3.10) is a weaker notion of solution than the classical Filippov formulation (see Remark 3.3). In Example 3.1 above, this means $b(0, x_2)$ is taken to be the smallest convex set containing all limit points of $b(y)$ as $y \rightarrow (0, x_2)$, that is, the line segment connecting $(-1, -1)$ and $(1, 1)$. The unique such solution is then the flow $\phi^{(0)}$. However, we can construct an example where even the Filippov flow is not unique.

Example 3.2. Taking $d = 2$ again, we set

$$b(x_1, x_2) = \begin{cases} (1, 1), & x_1 > 0 \text{ and } x_2 > 0, \\ (1, 0), & x_1 < 0 \text{ and } x_2 > 0, \\ (0, 1), & x_1 > 0 \text{ and } x_2 < 0, \\ (-1, -1), & x_1 < 0 \text{ and } x_2 < 0. \end{cases}$$

There are then three possible solution flows, even with the general Filippov formulation, when the starting point (x_1, x_2) satisfies $x_1 = x_2$ and $\tau > |x_1| = |x_2|$. Indeed, the nonuniqueness occurs once trajectories reach the origin. Then (3.10) admits the three solutions

$$\phi_{-\tau}^{(1)}(0) = -\tau(1, 0), \quad \phi_{-\tau}^{(2)}(0) = -\tau(0, 1), \quad \text{and} \quad \phi_{-\tau}^{(3)}(0) = 0 \quad \text{for } \tau > 0.$$

In fact, $\phi_{-\tau}^{(k)}$ is not continuous for any $\tau > 0$ and $k = 1, 2, 3$. This can be seen by considering the flows starting from $(x_1, x_2) < (0, 0)$ with $x_1 \neq x_2$: for $\tau > 0$,

$$\phi_{-\tau}^{(1)}(x) = \phi_{-\tau}^{(2)}(x) = \phi_{-\tau}^{(3)}(x) = \begin{cases} (x_1 + \tau, x_2 + \tau) & \text{if } 0 < \tau < -\max(x_1, x_2), \\ (x_1 - 2x_2 - \tau, 0) & \text{if } x_1 < x_2 < 0 \text{ and } \tau > |x_2|, \\ (0, x_2 - 2x_1 - \tau) & \text{if } x_2 < x_1 < 0 \text{ and } \tau > |x_1|. \end{cases}$$

In particular, given $\tau > 0$ and $\varepsilon > 0$, for $k = 1, 2, 3$,

$$\phi_{-\tau}^{(k)}(-\tau/2 + \varepsilon, -\tau/2) - \phi_{-\tau}^{(k)}(-\tau/2, -\tau/2 + \varepsilon) = \left(\frac{\tau}{2} + 2\varepsilon\right)(1, -1).$$

Poupaud and Rascle [55] explore the connection between the uniqueness (for every $x \in \mathbb{R}^d$) of Filippov solutions of (3.8) and a stable notion of measure-valued solutions to the continuity equation³

$$\partial_t f - \operatorname{div}(b(t, x)f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d \quad \text{and} \quad f(0, \cdot) = f_0. \quad (3.11)$$

As Example 3.2 shows, we cannot take this approach in analyzing (3.11), because there are three distinct measure-valued solutions when $f_0 = \delta_0$.

³Observe, from the minus sign in front of the velocity field, that this is not the same initial value problem as (4.2).

3.5. Stochastic differential equations

For b satisfying (3.2), we now consider the stochastic flow for the SDE

$$d_t \Phi_{t,s}(x) = b(t, \Phi_{t,s}(x)) dt + \sigma(t, \Phi_{t,s}(x)) dW_t, \quad t \in [s, T], \quad \Phi_{s,s}(x) = x. \quad (3.12)$$

Here, for some $m \in \mathbb{N}$, $W: [0, T] \times \Omega \rightarrow \mathbb{R}^m$ is an m -dimensional Wiener process on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and, for $(t, x) \in [0, T] \times \mathbb{R}^d$, σ is a $(d \times m)$ -dimensional matrix denoted coordinate-wise by σ_{ij} , $i = 1, 2, \dots, d$, $j = 1, 2, \dots, m$.

Because σ may be degenerate, the addition of noise does not necessarily imply the existence and uniqueness of a strong solution. In reproducing the theory for the ODE (3.1), we are therefore led to the condition

$$\begin{cases} \sigma \in L^2([0, T], C^{0,1}(\mathbb{R}^d; \mathbb{R}^{d \times m})) \text{ and,} \\ \text{for a.e. } t \in [0, T] \text{ and all } i = 1, 2, \dots, d, \quad j = 1, 2, \dots, m, \\ \sigma_{ij}(t, \cdot) \text{ is independent of } x_k \text{ for } k \neq i. \end{cases} \quad (3.13)$$

We assume in addition to (3.2) and (3.13) a bounded oscillation condition for b :

$$\begin{cases} \text{for some } C_2 \in L^1_+([0, T]), \text{ and for a.e. } (t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d, \\ |b(t, x) - b(t, y)| \leq C_2(t)(|x - y| + 1). \end{cases} \quad (3.14)$$

While not strictly necessary, this helps simplify some of the arguments by allowing certain regularizations of b to be globally Lipschitz.

The reformulation of (3.12) as a differential inclusion then takes the form

$$\begin{cases} d_t \Phi_{t,s}(x) = \alpha_t dt + \sigma(t, \Phi_{t,s}(x)) dW_t, \\ \alpha_t \in [b_*(t, \cdot), b^*(t, \cdot)](\Phi_{t,s}(x)) \\ \Phi_{s,s} = x. \end{cases} \quad \text{for a.e. } t \in [s, T], \quad (3.15)$$

Theorem 3.5. Assume b satisfies (3.2) and (3.14), and σ satisfies (3.13). Then the following hold with probability 1, for all $s \in [0, T]$:

- (a) There exist solutions of $\Phi_{\cdot,s}^+$ and $\Phi_{\cdot,s}^-$ of (3.15) such that, for all $t \in [s, T]$, $\Phi_{t,s}^+$ is upper semicontinuous and increasing and $\Phi_{t,s}^-$ is lower semicontinuous and increasing.
- (b) Any other solution Φ of (3.15) satisfies $\Phi_{\cdot,s}^- \leq \Phi_{\cdot,s} \leq \Phi_{\cdot,s}^+$.
- (c) If $0 \leq r \leq s \leq t \leq T$, then $\Phi_{t,s}^+ \circ \Phi_{s,r}^+ = \Phi_{t,r}^+$ and $\Phi_{t,s}^- \circ \Phi_{s,r}^- = \Phi_{t,r}^-$.
- (d) For all $t \in [s, T]$, $(\Phi_{t,s}^+)^* = \Phi_{t,s}^-$.
- (e) For a.e. $x \in \mathbb{R}^d$, (3.12) admits a unique integral solution, which is equal a.e. to $\Phi_{\cdot,s}^+$ and $\Phi_{\cdot,s}^-$.
- (f) There exists $\omega: [0, \infty) \rightarrow [0, \infty)$ such that $\lim_{r \rightarrow 0^+} \omega(r) = 0$ and, for all $f \in L^1(\mathbb{R}^d)$ and $0 \leq s \leq t \leq T$,

$$\mathbb{E}[\|f \circ \Phi_{t,s}\|_{L^1}] \leq e^{\omega(t-s)} \|f\|_{L^1}.$$

We first establish an analogue of Lemma 3.2.

Lemma 3.3. *Assume $B: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfies*

$$\begin{cases} \{t \mapsto \|\nabla B(t, \cdot)\|_\infty\} \in L^1([0, T]) \text{ and} \\ \partial_{x_i} B^j \geq 0 \text{ for all } i \neq j, \end{cases}$$

and σ satisfies (3.13). Let $X, Y, \alpha, \beta: [0, T] \rightarrow \mathbb{R}^d$ be adapted with respect to the Wiener process W such that, with probability 1, X and Y are continuous, $\alpha, \beta \in L^1([0, T])$,

$$\alpha_t \leq B(t, X_t) \quad \text{and} \quad \beta_t \geq B(t, Y_t) \quad \text{a.e. in } [0, T],$$

and

$$dX_t = \alpha_t dt + \sigma(t, X_t) dW_t, \quad dY_t = \beta_t dt + \sigma(t, Y_t) dW_t \quad \text{in } [0, T], \quad X_0 \leq Y_0.$$

Then $X_t \leq Y_t$ for all $t \in [0, T]$.

Proof. For $t \in [0, T]$, define $\gamma_t = \alpha_t - \beta_t$, which satisfies $\gamma \in L^1$ and

$$\gamma_t^i \leq B^i(t, X_t) - B^i(t, Y_t) = \sum_{j=1}^d \mu_t^{ij} (X_t^j - Y_t^j)$$

for a.e. $t \in [0, T]$, where

$$\mu_t^{ij} = \int_0^1 \partial_j B^i(t, \theta X_t + (1 - \theta)Y_t) d\theta.$$

Note that $\mu^{ij} \in L^1$ for all $i, j \in \{1, 2, \dots, d\}$ and $\mu^{ij} \geq 0$ for $i \neq j$. Define also, for $i = 1, 2, \dots, d$ and $k = 1, 2, \dots, m$,

$$v_t^{ik} = \int_0^1 \partial_{x_k} \sigma_{ik}(t, \theta X_t^i + (1 - \theta)Y_t^i) d\theta,$$

which satisfies $v^{ik} \in L^2$ (recall that σ_{ik} depends only on (t, x_i)). Therefore, if, for $i = 1, 2, \dots, d$, setting $\Delta^i = X^i - Y^i$, we find that, for $i = 1, 2, \dots, d$,

$$d\Delta_t^i = \gamma_t^i dt + \sum_{k=1}^m v_t^{ik} \Delta_t^i dW_t^k, \quad \gamma_t^i \leq \sum_{j=1}^d \mu_t^{ij} \Delta_t^j, \quad \Delta_0^i \leq 0.$$

Now, for $i = 1, 2, \dots, d$, let Z^i be the solution of the SDE

$$dZ_t^i = \left(\sum_{k=1}^m (v_t^{ik})^2 - \mu_t^{ii} \right) Z_t^i dt - \sum_{k=1}^m v_t^{ik} Z_t^i dW_t^k, \quad Z_0^i = 1.$$

Then

$$Z_t^i = \exp\left(\int_0^t \left(\frac{1}{2} \sum_{k=1}^m (v_s^{ik})^2 - \mu_s^{ii}\right) ds - \int_0^t \sum_{k=1}^m v_s^{ik} dW_s^k\right) \geq 0,$$

and so Itô's formula yields

$$d(Z_t^i \Delta_t^i) = (Z_t^i \gamma_t^i - \mu_t^{ii} Z_t^i \Delta_t^i) dt \leq \sum_{j \neq i} \mu_t^{ij} Z_t^i \Delta_t^j dt.$$

For some $\delta > 0$ and an adapted, absolutely continuous, and increasing $\omega: [0, T] \rightarrow \mathbb{R}_+$ with $\omega(0) = 1$, define

$$\tau := \inf\{t \in [0, T] : Z_t^i \Delta_t^i - \delta e^{\omega(t)} > 0 \text{ for some } i = 1, 2, \dots, d\}.$$

We have $\tau > 0$ because $Z_0^i \Delta_0^i \leq 0$. Assume for the sake of contradiction that $\tau < \infty$. If $i = 1, 2, \dots, m$ is such that $Z_\tau^i \Delta_\tau^i = \delta e^{\omega(\tau)}$, then

$$\delta e^{\omega(\tau)} = Z_\tau^i \Delta_\tau^i \leq \sum_{j \neq i} \int_0^\tau \mu_t^{ij} Z_t^i \Delta_t^j dt \leq \delta \sum_{j \neq i} \int_0^\tau \mu_t^{ij} R_t^{ij} e^{\omega(t)} dt,$$

where

$$R_t^{ij} = \frac{Z_t^i}{Z_t^j} = \prod_{k=1}^m \exp\left(\frac{1}{2} \int_0^t [(v_s^{ik})^2 - (v_s^{jk})^2 - \mu_s^{ii} + \mu_s^{jj}] ds + \int_0^t (v_s^{jk} - v_s^{ik}) dW_s^k\right).$$

If we choose ω to satisfy

$$\omega(0) = 1, \quad \omega'(t) = 2 \max_{i=1,2,\dots,d} \sum_{j \neq i} \mu_t^{ij} R_t^{ij} \geq 0,$$

then

$$\delta e^{\omega(\tau)} \leq \frac{\delta}{2} \int_0^\tau \omega'(t) e^{\omega(t)} dt = \frac{\delta}{2} (e^{\omega(\tau)} - 1),$$

which is a contradiction for any $\delta > 0$. This implies $\Delta_t^i \leq \delta e^{\omega(t)} (Z_t^i)^{-1}$ for all $i = 1, 2, \dots, m, t \in [0, T]$, and $\delta > 0$, and we conclude upon sending $\delta \rightarrow 0$. ■

Proof of Theorem 3.5. Let b^ε and b_ε be the sup- and inf-convolutions of b as in the proof of Proposition 3.1, and define the stochastic flows $\Phi^{+, \varepsilon}$ and $\Phi^{-, \varepsilon}$ by

$$d_t \Phi_{t,s}^{+, \varepsilon}(x) = b^\varepsilon(t, \Phi_{t,s}^{+, \varepsilon}(x)) dt + \sigma(t, \Phi_{t,s}^{+, \varepsilon}(x)) dW_t \quad \text{in } [s, T], \quad \Phi_{s,s}^{+, \varepsilon} = \text{Id}$$

and

$$d_t \Phi_{t,s}^{-, \varepsilon}(x) = b_\varepsilon(t, \Phi_{t,s}^{-, \varepsilon}(x)) dt + \sigma(t, \Phi_{t,s}^{-, \varepsilon}(x)) dW_t \quad \text{in } [s, T], \quad \Phi_{s,s}^{-, \varepsilon} = \text{Id}.$$

Parts (a), (b), and (c) are then established exactly as in the proof of Proposition 3.1, making use of the comparison result of Lemma 3.3 above.

Part (d), and then, as a consequence, (e), are proved similarly to the proof of Proposition 3.2 and Theorem 3.1, instead now using the one-sided mollifiers ρ^ε and ρ_ε to regularize b . Note that, in view of assumption (3.14), for some $C \in L^1_+([0, T])$,

$$|\nabla(b(t, \cdot) * \rho^\varepsilon)(x)| \leq \int |b(t, x) - b(t, y)| |\nabla \rho^\varepsilon(x - y)| dy \leq \frac{C(t)}{\varepsilon},$$

and similarly for $b * \rho_\varepsilon$. This allows for the use of the comparison Lemma 3.3 in the argument, which requires global Lipschitz regularity in view of the lack of finite speed of propagation.

Finally, (f) follows exactly as in the proof of Theorem 3.1. ■

4. Linear transport equations with increasing drift

For velocity fields b satisfying (3.2), we discuss the terminal value problem for the non-conservative equation

$$-\partial_t u - b(t, x) \cdot \nabla u = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T, \quad (4.1)$$

as well the initial value problem for the conservative equation

$$\partial_t f + \operatorname{div}(b(t, x)f) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad f(0, \cdot) = f_0. \quad (4.2)$$

The lower bound on $\operatorname{div} b$ implied by (3.2), which gave rise to the regular Lagrangian property for the unique flow ϕ in the previous section, here allows for a theory of weak solutions of (4.1) and (4.2) in L^p -spaces, due to the expansive property of the flow.

4.1. The nonconservative equation

It is important to note that solutions u of (4.1) taking values in L^p_{loc} cannot be understood in the sense of distributions. This is because, under assumption (3.2), $\operatorname{div} b$ is a measure that need not necessarily be absolutely continuous with respect to Lebesgue measure. Instead, we identify unique solution operators for (4.1) that are continuous on Lebesgue spaces and stable under regularizations. This is done through the relationship with the flow in the previous section, and also by characterizing the solutions using “one-sided” regularizations for increasing or decreasing solutions.

4.1.1. Representation formula. When b and \bar{u} are smooth and $t \in [0, T]$, the unique solution of the transport equation (4.1) in $[0, t] \times \mathbb{R}^d$ with terminal value \bar{u} at time t is

$$u(s, x) = S(s, t)\bar{u}(x) := \bar{u}(\phi_{t,s}(x)), \quad s \in [0, t], \quad (4.3)$$

where ϕ is the flow corresponding to the ODE (3.1). In view of Theorem 3.1, this formula extends to $\bar{u} \in L^p_{\text{loc}}$, $1 \leq p \leq \infty$ if the assumption on b is relaxed to (3.2). We thus identify a family of solution operators for (4.1) that are continuous on L^p and evolve continuously in time.

Theorem 4.1. Assume that b satisfies (3.2), and define the solution operators in (4.3) using the flow constructed in Section 3. Then the following hold:

- (a) For all $0 \leq s \leq t \leq T$, $S(s, t)$ is continuous on $L^p_{\text{loc}}(\mathbb{R}^d)$ for $1 \leq p \leq \infty$, and there exists $C > 0$ depending only on T such that, for any $R > 0$,

$$\|S(s, t)\bar{u}\|_{L^p(B_R)} \leq e^{d(\omega_1(t) - \omega_1(s))} \|\bar{u}\|_{L^p(B_{R+C(1+R)(\omega_0(t) - \omega_0(s))})}.$$

- (b) For all $0 \leq r \leq s \leq t \leq T$, we have $S(r, s) \circ S(s, t) = S(r, t)$.
(c) If $0 < t \leq T$, then $S(\cdot, t)\bar{u} \in C([0, t], L^p_{\text{loc}}(\mathbb{R}^d))$ if $\bar{u} \in L^p_{\text{loc}}$ for $p < \infty$, and $S(\cdot, t)\bar{u} \in C([0, t], L^\infty_{\text{loc}, w-\star}(\mathbb{R}^d))$ if $\bar{u} \in L^\infty_{\text{loc}}$.
(d) If $(b^\varepsilon)_{\varepsilon>0}$ is any family satisfying (3.2) uniformly in ε such that

$$\int_0^T \|\nabla b^\varepsilon(t, \cdot)\|_\infty dt < \infty \quad \text{for all } \varepsilon > 0$$

and $b^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} b$ a.e., $(u^\varepsilon_T)_{\varepsilon>0}$ are smooth functions such that $u^\varepsilon_T \xrightarrow{\varepsilon \rightarrow 0} u_T$ in L^p_{loc} for some $p < \infty$, and u^ε is the corresponding solution of (4.1), then, as $\varepsilon \rightarrow 0$, u^ε converges to $S(t, T)u_T$ strongly in $C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$. The same statement is true if u^ε solves

$$-\partial_t u^\varepsilon - b(t, x) \cdot \nabla u^\varepsilon = \varepsilon \Delta u^\varepsilon \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u^\varepsilon(T, \cdot) = u^\varepsilon_T. \quad (4.4)$$

Proof. Properties (a) and (b) follow immediately from Theorem 3.1, while properties (c) and (d) follow from Theorem 3.2 and the dominated convergence theorem (and see also Theorem 3.4). For the statement involving the viscous equation (4.4), we remark that, in that case, u^ε is given by $u^\varepsilon(t, x) = u^\varepsilon_T(\phi^\varepsilon_{T,t}(x))$, where ϕ^ε is the stochastic flow corresponding to (3.7). The proof is then finished because of Theorem 3.3. ■

Remark 4.1. The uniqueness of the semigroup is a consequence of the uniqueness of the flow established in the previous section. Note, however, that, solely under the assumption $\text{div } b \geq -C_1$ for some $C_1 \in L^1_+([0, T])$, any weakly limiting family of solution operators must lead to solutions in $C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$, in the strong topology.

More precisely, for $\varepsilon > 0$, we have the easy a priori bound

$$\|S^\varepsilon(s, t)\bar{u}\|_{L^p} \leq \exp\left(\int_s^t C_1(r) dr\right) \|\bar{u}\|_{L^p}, \quad \bar{u} \in L^p.$$

It follows from a diagonal argument that there exist $\varepsilon_n \rightarrow 0$ and a family of continuous linear operators on L^p such that $S(r, s)S(s, t) = S(r, t)$ for $r \leq s \leq t$, and, for all $t \in [0, T]$,

$$S^{\varepsilon_n}(\cdot, t)\bar{u} \rightharpoonup S(\cdot, t)\bar{u} \quad \text{weakly in } L^\infty([0, T], L^p(\mathbb{R}^d)) \text{ for all } \bar{u} \in L^p.$$

In particular, $u := S(\cdot, t)\bar{u} \in C([0, T], L^p_w(\mathbb{R}^d))$, and so, for any $s \in [0, t]$,

$$\liminf_{h \rightarrow 0^+} \|u(s+h, \cdot)\|_{L^p} \geq \|u(s, \cdot)\|.$$

On the other hand,

$$\|u(s+h, \cdot)\|_{L^p} \leq \exp\left(\int_s^{s+h} C_1(r) dr\right) \|u(s, \cdot)\|_{L^p},$$

so that

$$\limsup_{h \rightarrow 0^+} \|u(s+h, \cdot)\|_{L^p} \leq \|u(s, \cdot)\|.$$

We find then that $\|u(s+h, \cdot)\|_{L^p} \rightarrow \|u(s, \cdot)\|_{L^p}$, which, coupled with the weak convergence, means that $u(s+h, \cdot) \rightarrow u(s, \cdot)$ strongly in L^p if $p > 1$ (and so also in L^1_{loc}). A similar argument holds with h replaced by $-h$.

The following renormalization property for the solution operator $S(s, t)$, $0 \leq s \leq t \leq T$, is then immediate from the formula.

Proposition 4.1. *If $\beta: \mathbb{R} \rightarrow \mathbb{R}$ is smooth and $|\beta(s)| \leq C(1 + |s|^r)$ for some $r \geq 1$, then, for all $\tilde{u} \in L^r_{\text{loc}}$, $\beta(S(s, t)\tilde{u}) = S(s, t)(\beta \circ \tilde{u})$.*

4.1.2. Characterizing increasing/decreasing solutions. The solution of the transport equation in the previous subsection was characterized as the unique limit under arbitrary regularizations, as well as through the formula involving the regular Lagrangian flow. The remainder of the section is dedicated to understanding further ways to characterize the solution, and in particular on the level of equation (4.1) itself. This will become useful in the study of nonlinear equations in the final section.

We first observe that, if u_T is increasing/decreasing with respect to the partial order (2.1), then so is $u(t, \cdot)$ for all $t \in [0, T]$. While this is immediately clear from the formula $u(t, \cdot) = u_T \circ \phi_{T,t}$ and the fact that $\phi_{T,t}$ is increasing, it can also be seen directly from the equation. Indeed, if u is a smooth solution of (4.1) and $v_i = \partial_{x_i} u$, $i = 1, 2, \dots, d$, then

$$-\partial_t v_i - b \cdot \nabla v_i - (\partial_{x_i} b^j) v_i = \sum_{j \neq i} (\partial_{x_i} b^j) v_j.$$

Therefore, if $v_i \geq 0$ (or $v_i \leq 0$) when $t = T$ for all $i = 1, 2, \dots, d$, then the same is true for $t < T$ by the maximum principle Lemma 2.6. The result for general b satisfying (3.2) follows from approximating b and using the limiting result in Theorem 4.1. By linearity, we have thus established the following:

Proposition 4.2. *For all $0 \leq s \leq t \leq T$, $S(s, t): \text{ABV} \rightarrow \text{ABV}$.*

In particular, since ABV is densely contained in $L^p_{\text{loc}}(\mathbb{R}^d)$ and $S(s, t)$ is continuous on L^p_{loc} , Proposition 4.2 implies that belonging to ABV is a suitable criterion for the propagation of compactness. Notice for instance that this provides another proof of the fact that the convergence statements in Theorem 4.1 are with respect to *strong* convergence in $C([0, T], L^p_{\text{loc}})$ for $p < \infty$.

We now demonstrate how the propagation of the increasing or decreasing property leads to a method for characterizing solutions of (4.1), independently of the solution formula. The idea is to regularize u in a one-sided manner, as in Section 2.3.

Formally, if u solves (4.1) and, for $(t, x) \in [0, T] \times \mathbb{R}^d$ (recall that ω_1 is as in (3.3)),

$$\tilde{u}(t, x) = u(t, e^{\omega_1(T)-\omega_1(t)}x)$$

and

$$\tilde{b}(t, x) = e^{-(\omega_1(T)-\omega_1(t))}b(t, e^{\omega_1(T)-\omega_1(t)}x) + C_1(t)x,$$

then \tilde{u} solves (4.1) with b replaced by \tilde{b} , and \tilde{b} satisfies (3.2) with $C_1 = 0$ and a possibly different C_0 . We may therefore assume here, without loss of generality, that $b(t, \cdot)$ is increasing for a.e. $t \in [0, T]$.

Recall the definition of the one-sided mollifiers ρ_ε and ρ^ε defined in (2.4).

Definition 4.1. A function $u: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is called a supersolution (resp. subsolution) of (4.1) if, for all $t \in [0, T]$, $u(t, \cdot)$ is decreasing, and, if $u_\varepsilon = u * \rho_\varepsilon$ and $u^\varepsilon = u * \rho^\varepsilon$, then

$$-\partial_t u_\varepsilon - b(t, x) \cdot \nabla u_\varepsilon \geq 0 \quad (\text{resp. } -\partial_t u^\varepsilon - b(t, x) \cdot \nabla u^\varepsilon \leq 0) \quad (4.5)$$

for a.e. $(t, x) \in [0, T] \times \mathbb{R}^d$.

A solution is both a sub- and a supersolution.

Remark 4.2. A well-posed notion of sub- and supersolutions can be defined where u is approximated using, instead of the one-sided mollifiers, the method of inf- and sup-convolution:

$$C_\varepsilon[u(t, \cdot)](x) = \inf_y \left\{ u(t, x - y) + \frac{|y|}{\varepsilon} \right\} \quad \text{and} \quad C^\varepsilon[u(t, \cdot)](x) = \sup_y \left\{ u(t, x - y) - \frac{|y|}{\varepsilon} \right\}.$$

Theorem 4.2. Assume b satisfies (3.2) with $C_1 = 0$. Then, for all $R > 0$, there exists a modulus of continuity $\omega_R: [0, \infty) \rightarrow [0, \infty)$ such that, if $u, v: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ are respectively a sub- and a supersolution of (4.1) in the sense of Definition 4.1, then, for all $t \in [0, T]$ and $p \in [1, \infty)$,

$$\int_{B_R} (u(t, x) - v(t, x))_+^p dx \leq \int_{B_{R+\omega_R(T-t)}} (u(T, x) - v(T, x))_+^p dx. \quad (4.6)$$

In particular, for any decreasing $u_T: \mathbb{R}^d \rightarrow \mathbb{R}^d$, there exists a unique solution u of (4.1) in the sense of Definition 4.1, which is given by $u(t, \cdot) = S(t, T)u_T$, and which is continuous a.e. in $[0, T] \times \mathbb{R}^d$.

Proof. Set $u^\varepsilon = u * \rho^\varepsilon$ and $v_\varepsilon = v * \rho_\varepsilon$; then, combining the inequalities for u^ε and v_ε given by Definition 4.1, we obtain

$$-\partial_t (u^\varepsilon - v_\varepsilon) - b(t, x) \cdot \nabla (u^\varepsilon - v_\varepsilon) \leq 0.$$

Let $\beta: \mathbb{R} \rightarrow \mathbb{R}$ be smooth and increasing. Multiplying the above inequality by the positive term $\beta'(u^\varepsilon - v_\varepsilon)$ yields

$$-\partial_t \beta(u^\varepsilon - v_\varepsilon) - b(t, x) \cdot \nabla \beta(u^\varepsilon - v_\varepsilon) \leq 0.$$

We may then take $\beta(r) = r_+^p$ (arguing with an extra layer of regularizations if $p = 1$) and find that, in the sense of distributions,

$$-\partial_t(u^\varepsilon - v_\varepsilon)_+^p - b(t, x) \cdot \nabla(u^\varepsilon - v_\varepsilon)_+^p \leq 0.$$

Let $\hat{\psi}: [0, \infty) \rightarrow [0, \infty)$ be smooth and decreasing, such that $\hat{\psi} = 1$ on $[0, 1]$ and $\hat{\psi} = 0$ on $[2, \infty)$, and set $\psi(t, x) := \hat{\psi}(|x|/R(t))$ for some $R: [0, T] \rightarrow [0, \infty)$ to be determined. Using ψ as a test function, we discover

$$\begin{aligned} & -\partial_t \int_{\mathbb{R}^d} (u^\varepsilon(t, x) - v_\varepsilon(t, x))_+^p \psi(t, x) dx \\ & \leq \langle -\partial_t \psi(t, \cdot) - \operatorname{div}(b(t, \cdot) \psi(t, \cdot)), (u^\varepsilon(t, \cdot) - v_\varepsilon(t, \cdot))_+^p \rangle \\ & \leq \int_{\mathbb{R}^d} (u^\varepsilon(t, x) - v_\varepsilon(t, x))_+^p (-\partial_t \psi(t, x) - b(t, x) \cdot \nabla \psi(t, x)) dx, \end{aligned}$$

where in the last line we used the fact that $\operatorname{div} b \geq 0$ and $\psi \geq 0$. Using the fact that $\hat{\psi}' \leq 0$, with $\hat{\psi}' \neq 0$ only if $r \in [1, 2]$, we find that

$$\begin{aligned} -\partial_t \psi - b \cdot \nabla \psi &= -\hat{\psi}'(x/R(t)) \left(\frac{\dot{R}(t)}{R(t)^2} - \frac{b(t, x) \cdot x}{R(t)|x|^2} \right) \\ &\geq -R(t)^{-1} \hat{\psi}'(x/R(t)) (\dot{R}(t) - 2C_0(t)(1 + 2R(t))). \end{aligned}$$

For $t_0 \in [0, T]$, this is made nonnegative on $[t_0, T]$ by choosing, for any fixed $R > 0$,

$$R(t) = R e^{4[\omega_0(t) - \omega_0(t_0)]} + \frac{1}{2} (e^{4[\omega_0(t) - \omega_0(t_0)]} - 1),$$

and so

$$\int_{\mathbb{R}^d} (u^\varepsilon(t_0, x) - v_\varepsilon(t_0, x))_+^p \hat{\psi} \left(\frac{|x|}{R(t_0)} \right) dx \leq \int_{\mathbb{R}^d} (u^\varepsilon(T, x) - v_\varepsilon(T, x))_+^p \hat{\psi} \left(\frac{|x|}{R(T)} \right) dx.$$

We may then choose functions $\hat{\psi}$ that approximate $\mathbf{1}_{[0,1]}$ from above, and then, by the monotone convergence theorem,

$$\int_{B_R} (u^\varepsilon(t_0, x) - v_\varepsilon(t_0, x))_+^p dx \leq \int_{B_{R(T)}} (u^\varepsilon(T, x) - v_\varepsilon(T, x))_+^p dx.$$

The proof of (4.6) is finished upon sending $\varepsilon \rightarrow 0$ and setting

$$\omega_R(r) := \sup_{0 \leq s \leq t \leq s+r \leq T} \left\{ \left(R + \frac{1}{2} \right) (e^{4[\omega_0(t) - \omega_0(s)]} - 1) \right\}.$$

The comparison inequality (4.6) implies uniqueness for a solution with terminal condition u_T , and so it remains to show that $S(t, T)u_T$ is a solution in the sense of Definition 4.1.

Let $(b^\delta)_{\delta>0}$ be a family of smooth functions satisfying (3.2) with $C_1 = 0$ and uniform C_0 and converging a.e. to b as $\delta \rightarrow 0$, and let $u_T^\delta = u_T * \gamma_\delta$ for a family of standard mollifiers $(\gamma_\delta)_{\delta>0}$. Note then that, for $\delta > 0$, u_T^δ is decreasing, and, as $\delta \rightarrow 0$, $u_T^\delta \rightarrow u_T$ in L^p_{loc} for all $p < \infty$. Let u^δ be the corresponding solution of the terminal value problem (4.1). It follows that $u^\delta(t, \cdot)$ is decreasing for all $t \in [0, T]$.

For any $\delta > 0$ and $\varepsilon > 0$,

$$\begin{aligned} & -\partial_t(u^\delta * \rho^\varepsilon) - b^\delta(t, x) \cdot \nabla(u^\delta * \rho^\varepsilon) \\ & = \int_{\mathbb{R}^d} (b^\delta(t, y) - b^\delta(t, x)) \cdot \nabla u^\delta(t, y) \rho^\varepsilon(x - y) dy \leq 0, \end{aligned}$$

where the last inequality follows from the fact that $\rho^\varepsilon(x - y) \neq 0$ only if $x \leq y$, $b^\delta(t, \cdot)$ is increasing, and $\nabla u^\delta \leq 0$. Sending $\delta \rightarrow 0$, it follows from Theorem 4.1 that, in the sense of distributions, if $u = S(\cdot, T)u_T$ and $u^\varepsilon = u * \rho^\varepsilon$,

$$-\partial_t u^\varepsilon - b \cdot \nabla u^\varepsilon \leq 0.$$

It follows that $S(\cdot, T)u_T$ is a subsolution. A similar argument shows that it is a supersolution, and therefore the unique solution in the sense of Definition 4.1.

Now, for some $M > 0$, define $\psi(t) = \exp(M\omega_1(t))$ and

$$\tilde{u}^\delta(t, x) := u^\delta(t, x + \psi(t)\mathbf{1}).$$

For any fixed $R > 0$, there exists $M > 0$ such that, in view of the linear growth of b^δ given by (3.2) uniformly in δ , and the fact that u^δ is decreasing in the spatial variable, for any $x \in [-R, R]^d$,

$$\begin{aligned} \frac{\partial}{\partial t} \tilde{u}^\delta(t, x) & = (-\nabla u^\delta) \cdot (b^\delta(t, x + \psi(t)\mathbf{1}) - \psi'(t)\mathbf{1}) \\ & \leq C_1(t)(-\nabla u^\delta) \cdot ((1 + |x| + \psi(t)) - M\psi(t)) \leq 0. \end{aligned}$$

Since M is independent of δ , we may send $\delta \rightarrow 0$ and conclude that $\tilde{u}(t, x) := u(t, x + \psi(t))$ is increasing on $[0, T] \times [-R, R]^d$, and therefore continuous a.e. in that set by Lemma 2.3. The transformation leading from u to \tilde{u} preserves null sets, and we conclude that u is continuous almost everywhere. ■

Remark 4.3. The idea behind Definition 4.1 is to establish a sign for the commutator between convolution and differentiation along irregular vector fields, as compared to the work of DiPerna and the first author [30], where the commutator is shown to be small for Sobolev vector fields. We must take convolution kernels with a specific one-sided structure in order to analyze the commutators; a less crude example of this idea is seen in the work of Ambrosio [2] for general BV velocity fields.

A different notion of sub- and supersolutions, which also selects $S(\cdot, T)u_T$ as the unique solution of (4.1), can be obtained by instead regularizing b . Recall that we have

assumed without loss of generality that b is increasing. Then a decreasing function u can be said to be a subsolution (resp. supersolution) if, for all $\varepsilon > 0$, in the sense of distributions,

$$-\partial_t u - b_\varepsilon \cdot \nabla u \leq 0 \quad (\text{resp. } -\partial_t u - b^\varepsilon \cdot \nabla u \geq 0),$$

where $b_\varepsilon \leq b \leq b^\varepsilon$ are one-sided regularizations of b , for example the one-sided mollifiers or the inf- and sup-convolutions. The notion of sub- and supersolutions in Definition 4.1 turns out to be more amenable to the study of the nonlinear systems in Section 5.

Remark 4.4. Throughout this section, we have studied the setting where u_T , and therefore $u(t, \cdot)$ for $t < T$, is decreasing. The same analysis can be achieved for increasing solutions, in which case the inequalities in (4.5) are reversed. Note, in particular, that the solution flows $\phi_{t,s}(x)$ are vector-valued solutions in this sense, that is,

$$-\partial_s \phi_{t,s}(x) - b(s, x) \cdot \nabla \phi_{t,s}(x) = 0, \quad \phi_{t,t}(x) = x.$$

We now observe that the full family of solution operators $S(s, t): L^p_{\text{loc}} \rightarrow L^p_{\text{loc}}$, $0 \leq s \leq t \leq T$, can be constructed independently of the ODE flows $\phi_{s,t}$, which can then be recovered with the theory of renormalization. Indeed, Definition 4.1, and its counterpart for increasing solutions, can be used to define $S(s, t)\bar{u}$ for any $\bar{u} \in \text{ABV}$. The density of ABV in $L^p_{\text{loc}}(\mathbb{R}^d)$, and the L^p -continuity of $S(s, t)$, then allow the solution operators to be continuously extended to L^p_{loc} . It is, however, not clear whether solutions of (4.1) can be characterized for arbitrary $u_T \in L^p_{\text{loc}}$, other than by formula (4.3) or as limits of solutions to regularized equations.

4.1.3. Lower-order terms. We briefly explain how to extend the above results to equations with additional lower-order terms, as in

$$-\partial_t u - b(t, x) \cdot \nabla u - c(t, x)u - d(t, x) = 0 \quad \text{in } [0, T] \times \mathbb{R}^d, \quad u(T, \cdot) = u_T, \quad (4.7)$$

for functions

$$c \in L^1([0, T], L^\infty_{\text{loc}}) \quad \text{and} \quad d \in L^1([0, T], L^q_{\text{loc}}) \quad \text{for some } q \in [1, \infty]. \quad (4.8)$$

Theorem 4.3. Assume $p \in [1, \infty) \cap [1, q]$ and $u_T \in L^p_{\text{loc}}$. Then there exists a unique function $u \in C([0, T], L^p_{\text{loc}})$ with the following properties:

- (a) There exist $C_2 \in L^1_+([0, T])$, and, for $R > 0$, a modulus $\omega_R: [0, T] \rightarrow [0, T]$, depending only on the assumptions in (3.2) and (4.8) such that, for $R > 0$,

$$\|u(t, \cdot)\|_{L^p(B_R)} \leq \exp\left(\int_t^T C_2(s) ds\right) \|u_T\|_{L^p_{B_{R+\omega_R(T-t)}}} + \int_t^T C_2(s) ds.$$

- (b) Let $(b^\varepsilon)_{\varepsilon>0}$, $(c^\varepsilon)_{\varepsilon>0}$, and $(d^\varepsilon)_{\varepsilon>0}$ be families satisfying (3.2) and (4.8) uniformly in ε , such that, as $\varepsilon \rightarrow 0$, $(b^\varepsilon, c^\varepsilon, d^\varepsilon) \rightarrow (b, c, d)$ a.e. Let $(u^\varepsilon_T)_{\varepsilon>0}$ be a family of smooth functions approximating u_T in L^p_{loc} , and let u^ε be the corresponding

solution of (4.7). Then, as $\varepsilon \rightarrow 0$, u^ε converges strongly to u in $C([0, T], L_{\text{loc}}^p)$. The same statement is true if u^ε solves

$$\begin{cases} -\partial_t u^\varepsilon - b(t, x) \cdot \nabla u^\varepsilon - c(t, x) u^\varepsilon - d(t, x) = \varepsilon \Delta u^\varepsilon & \text{in } [0, T] \times \mathbb{R}^d, \\ u^\varepsilon(T, \cdot) = u_T^\varepsilon. \end{cases}$$

(c) For $(t, x) \in [0, T] \times \mathbb{R}^d$, u has the formula

$$\begin{aligned} u(t, x) = & u_T(\phi_{T,t}(x)) \exp\left(\int_t^T c(s, \phi_{s,t}(x)) ds\right) \\ & + \int_t^T d(s, \phi_{s,t}(x)) \exp\left(\int_t^s c(r, \phi_{r,t}(x)) dr\right) ds. \end{aligned} \quad (4.9)$$

Analogous statements hold when $p = q = \infty$, in which case $u \in C([0, T], L_{\text{loc}, w-\star}^\infty)$ and the convergence in part (b) is weak- \star in L^∞ .

Proof. For b^ε , c^ε , d^ε , u_T^ε , and u^ε as in the statement of part (b), if $(\phi_{t,s}^\varepsilon(x))_{s,t \in [0,T]}$ denotes the corresponding smooth flow, we have the formula

$$\begin{aligned} u^\varepsilon(t, x) = & u_T^\varepsilon(\phi_{T,t}^\varepsilon(x)) \exp\left(\int_t^T c^\varepsilon(s, \phi_{s,t}^\varepsilon(x)) ds\right) \\ & + \int_t^T d^\varepsilon(s, \phi_{s,t}^\varepsilon(x)) \exp\left(\int_t^s c^\varepsilon(r, \phi_{r,t}^\varepsilon(x)) dr\right) ds. \end{aligned}$$

The convergence statements, and thus the formula in part (c), are then proved just as in Theorem 3.2. In particular, arguing just as in that proof, we have

$$\lim_{\varepsilon \rightarrow 0} (\|u_T^\varepsilon(\cdot, \phi_{\cdot,t}^\varepsilon) - u_T(\cdot, \phi_{\cdot,t})\|_{L^1([t,T], L_{\text{loc}}^p)} + \|d^\varepsilon(\cdot, \phi_{\cdot,t}^\varepsilon) - d(\cdot, \phi_{\cdot,t})\|_{L^1([t,T], L_{\text{loc}}^p)}) = 0$$

and, for all $r < \infty$,

$$\lim_{\varepsilon \rightarrow 0} \|c^\varepsilon(\cdot, \phi_{\cdot,t}^\varepsilon) - c(\cdot, \phi_{\cdot,t})\|_{L^1([t,T], L_{\text{loc}}^r)} = 0.$$

In addition, $c^\varepsilon(t, \phi_{t,s}^\varepsilon)$ is uniformly bounded in $L^1([0, T], L_{\text{loc}}^\infty)$, in view of Lemma 3.1, and so we conclude parts (b) and (c) by the dominated convergence theorem. The L^p -estimates in part (a) are proved just as before, either from the regularized equation itself or from (4.9), using the lower bound on the divergence of b . ■

Sub- and supersolutions can be characterized when u_T is increasing/decreasing under the additional assumption that

$$c = c(t) \in L_+^1([0, T]) \quad \text{and} \quad d(t, \cdot) \text{ is decreasing for all } t \in [0, T]. \quad (4.10)$$

The propagation result Proposition 4.2 is then easily generalized (the proof is almost identical and so we omit it):

Proposition 4.3. Assume (3.2), (4.8), and (4.10). If $u_T \in L_{\text{loc}}^\infty$ is increasing (decreasing) and u is the solution of (4.7) specified by Theorem 4.3, then, for all $t < T$, $u(t, \cdot)$ is increasing (decreasing).

Definition 4.1 is then generalized as follows:

Definition 4.2. A function $u: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is called a supersolution (resp. subsolution) of (4.7) if, for all $t \in [0, T]$, $u(t, \cdot)$ is decreasing, and, if $u_\varepsilon = u * \rho_\varepsilon$ and $u^\varepsilon = u * \rho^\varepsilon$, then, for a.e. $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$\begin{aligned} -\partial_t u_\varepsilon - b(t, x) \cdot \nabla u_\varepsilon - c(t) u_\varepsilon - d(t, x) &\geq 0 \\ (\text{resp. } -\partial_t u^\varepsilon - b(t, x) \cdot \nabla u^\varepsilon - c(t) u^\varepsilon - d(t, x) &\leq 0). \end{aligned} \quad (4.11)$$

A solution is both a sub- and a supersolution.

Finally, the following is proved almost identically to Theorem 4.2:

Theorem 4.4. Assume b satisfies (3.2) with $C_1 = 0$, and c and d satisfy (4.8) and (4.10). Then, for all $R > 0$, there exist moduli of continuity $\omega, \omega_R: [0, \infty) \rightarrow [0, \infty)$ such that, if $u, v: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ are respectively a sub- and a supersolution of (4.7) in the sense of Definition 4.2, then, for all $t \in [0, T]$ and $p \in [1, \infty)$,

$$\int_{B_R} (u(t, x) - v(t, x))_+^p dx \leq e^{\omega(T-t)} \int_{B_{R+\omega_R(T-t)}} (u(T, x) - v(T, x))_+^p dx.$$

In particular, for any decreasing $u_T \in L_{\text{loc}}^p(\mathbb{R}^d)$, there exists a unique solution u of (4.7) in the sense of Definition 4.2, which is given by (4.9), and which is continuous a.e. in $[0, T] \times \mathbb{R}^d$.

4.2. The conservative equation

In contrast to the theory for the nonconservative equation, solutions of (4.2) belonging to Lebesgue spaces can be made sense of in the sense of distributions. However, under assumption (3.2), these are not in general unique, as the simple example $b(x) = \text{sgn } x$ on \mathbb{R} shows. Drawing an analogy once more with the setting studied in [12, 13, 48] of half-Lipschitz velocity fields, the “good” (stable) solution of (4.2) is identified using a particular solution formula, and, in [48], this is shown to coincide with the pushforward by the regular Lagrangian flow of the initial density. As discussed in Section 3.4 above, the former strategy is unavailable; however, in view of the theory of the nonconservative equation and the forward regular Lagrangian flow that was built in the previous section, we may define solutions by duality.

4.2.1. Duality solutions. For $0 \leq s \leq t \leq T$ denote by $S^*(t, s)$ the adjoint of the solution operator $S(s, t)$.

Theorem 4.5. Let $1 \leq p \leq \infty$ and $f_0 \in L_{\text{loc}}^p$, and define $f(t, x) = S^*(t, 0) f_0$. Then $f \in C([0, T], L_{\text{loc}}^p(\mathbb{R}^d))$ if $p < \infty$ or $f \in C([0, T], L_{\text{loc}, w-\star}^\infty)$ if $p = \infty$, and f is a distributional

solution of (4.2) and

$$S^*(t, 0) f_0 = (\phi_{t,0})_{\#} f_0, \quad (4.12)$$

where ϕ is the flow obtained in Section 3. If $(b^\varepsilon)_{\varepsilon>0}$ is a family of smooth functions satisfying (3.2) uniformly in $\varepsilon > 0$ and converging a.e. to b as $\varepsilon \rightarrow 0$, $(f_0^\varepsilon)_{\varepsilon>0}$ is a family of approximations of f_0 in L^p_{loc} for $1 \leq p < \infty$, and f^ε is the corresponding solution of (4.2), then, as $\varepsilon \rightarrow 0$, $f^\varepsilon \rightarrow f$ weakly in $C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$. The same is true if f^ε is taken to be the unique smooth solution of

$$\partial_t f^\varepsilon + \operatorname{div}(b f^\varepsilon) = \varepsilon \Delta f^\varepsilon \quad \text{in } [0, T] \times \mathbb{R}^d, \quad f^\varepsilon(0, \cdot) = f_0^\varepsilon,$$

and analogous convergence statements hold in the weak- \star sense if $p = \infty$.

Proof. The identity (4.12) follows immediately from (4.3); observe that it is well defined for $f_0 \in L^p_{\text{loc}}$ in view of the regular Lagrange property (3.5).

We now prove the convergence statements, and it suffices to prove the results for $p < \infty$, since $L^\infty_{\text{loc}} \subset L^p_{\text{loc}}$ for any $p < \infty$. Let $(b^\varepsilon)_{\varepsilon>0}$, $(f_0^\varepsilon)_{\varepsilon>0}$, $(f^\varepsilon)_{\varepsilon>0}$ be as in the statement. In view of the lower bound on the divergence, it is straightforward to prove a priori L^p bounds. Namely, for all $R > 0$, there exists a modulus of continuity $\omega_R: [0, \infty) \rightarrow [0, \infty)$ depending only on T and C_0 from (3.2), such that, for all $\varepsilon > 0$ and $t \in [0, T]$,

$$\int_{B_R} |f^\varepsilon(t, x)|^p dx \leq e^{d(p-1)\omega_1(t)} \int_{B_{R+\omega_R(t)}} |f_0^\varepsilon(x)|^p dx. \quad (4.13)$$

It follows that there exist a subsequence $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$ and $f \in L^\infty([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$ such that $f^{\varepsilon_n} \rightarrow f$ weakly in $L^\infty([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$. Sending $n \rightarrow \infty$ for $\varepsilon = \varepsilon_n$ in the equation

$$\partial_t f^{\varepsilon_n} + \operatorname{div}(b^{\varepsilon_n} f^{\varepsilon_n}) = 0,$$

using the fact that b^ε converges strongly to b in L^p_{loc} for any $p < \infty$, we find that f is a distributional solution of (4.2), and thus, moreover, $f \in C([0, T], L^p_{\text{loc},w}(\mathbb{R}^d))$. The weak- \star convergence statements when $p = \infty$ are proved analogously.

Fix $\bar{u} \in C^1_c(\mathbb{R}^d)$ and $t_0 \in [0, T]$, and let u^ε denote the solution of (4.1) with terminal value $u^\varepsilon(t_0, \cdot) = \bar{u}$; in view of the results of the previous subsection, u^ε has compact support in $[0, T] \times \mathbb{R}^d$ uniformly in $\varepsilon > 0$.

Exploiting the duality of the equations and integrating by parts gives

$$\int_{\mathbb{R}^d} \bar{u}(x) f^\varepsilon(t_0, x) dx = \int_{\mathbb{R}^d} u^\varepsilon(0, x) f_0^\varepsilon(x) dx.$$

Sending $n \rightarrow \infty$ for $\varepsilon = \varepsilon_n$ gives the identity

$$\int_{\mathbb{R}^d} \bar{u}(x) f(t_0, x) dx = \int_{\mathbb{R}^d} u(0, x) f_0(x) dx.$$

It follows that $f(t_0, \cdot) = S(t_0, 0) f_0$, and we therefore have the full convergence for all $\varepsilon \rightarrow 0$. An identical argument can be used to prove the convergence statement for the vanishing viscosity limit.

Finally, the fact that f is continuous from $[0, T]$ into $L^p_{\text{loc}}(\mathbb{R}^d)$ with the strong topology is a consequence of the continuity of the upper bound in the L^p -estimate (4.13) (see Remark 4.1). ■

Remark 4.5. It is clear from the proof above that, when $1 < p < \infty$, the initial functions f_0^ε need only converge weakly in L^p_{loc} to f_0 as $\varepsilon \rightarrow 0$.

Remark 4.6. It is not clear whether the convergence of regularizations f^ε to the duality solution f of (4.2) can be upgraded to strong convergence, except when $d = 1$, in which case (3.2) coincides with a half-Lipschitz condition on b (see [48, Theorem 4.2]).

4.2.2. Nonnegative solutions and renormalization. It is by now standard in the theory of the continuity equation (4.2) that there is uniqueness of nonnegative solutions when the measure f_0 is concentrated on sets where the ODE (3.1) has a unique solution; see [2, 3]. In view of Theorem 3.1, we then have the following:

Theorem 4.6. Assume $1 \leq p \leq \infty$, $f_0 \in L^p_{\text{loc}}(\mathbb{R}^d)$, and $f_0 \geq 0$. Then $f \in C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$ is a nonnegative distributional solution of (4.2) if and only if $f := S^*(\cdot, 0)f_0$.

We present here an alternative proof using the characterization of f as the duality solution, in order to emphasize again that the theory in this section can be developed independently of the analysis of the ODE (3.1) in the previous section.

Proof of Theorem 4.6. By Theorem 4.5, $f = S^*(\cdot, 0)f_0$ is a distributional solution, and its nonnegativity can be seen through an approximation argument, since weak convergence preserves nonnegativity.

Assume now that $f \in C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$ is a nonnegative distributional solution of (4.2). Fix $t_0 \in [0, T]$ and a decreasing function $\bar{u}: \mathbb{R}^d \rightarrow \mathbb{R}$, and let $u = S(\cdot, t_0)\bar{u}$. Then $u(t, \cdot)$ is decreasing for $t \in [0, t_0]$, and is a solution of (4.1) in the sense of Definition 4.1. In particular, $u_\varepsilon = u * \rho_\varepsilon$ and $u^\varepsilon = u * \rho^\varepsilon$ satisfy, respectively,

$$-\partial_t u_\varepsilon - b \cdot \nabla u_\varepsilon \geq 0 \quad \text{and} \quad -\partial_t u^\varepsilon - b \cdot \nabla u^\varepsilon \leq 0 \quad \text{in } [0, t_0] \times \mathbb{R}^d.$$

Let $\psi \in C^1_c(\mathbb{R}^d)$ with $\psi \geq 0$. Using $u_\varepsilon(t, x)\psi(x)$ as a test function for f , we find that

$$\begin{aligned} & \int_{\mathbb{R}^d} f(t_0, x)(\bar{u} * \rho_\varepsilon)(x)\psi(x) dx - \int_{\mathbb{R}^d} f_0(x)u_\varepsilon(0, x)\psi(x) dx \\ &= \int_0^{t_0} \int_{\mathbb{R}^d} f(s, x) [\partial_s (u_\varepsilon(s, x)\psi(x)) + b(s, x) \cdot \nabla (u_\varepsilon(s, x)\psi(x))] dx ds \\ &\leq \int_0^{t_0} \int_{\mathbb{R}^d} f(s, x)u_\varepsilon(s, x)b(s, x) \cdot \nabla \psi(x) dx ds, \end{aligned}$$

where we have used the nonnegativity of ψ and f . Similarly,

$$\begin{aligned} & \int_{\mathbb{R}^d} f(t_0, x)(\bar{u} * \rho^\varepsilon)(x)\psi(x) dx - \int_{\mathbb{R}^d} f_0(x)u^\varepsilon(0, x)\psi(x) dx \\ &\geq \int_0^{t_0} \int_{\mathbb{R}^d} f(s, x)u^\varepsilon(s, x)b(s, x) \cdot \nabla \psi(x) dx ds. \end{aligned}$$

Sending $\varepsilon \rightarrow 0$, we conclude that

$$\begin{aligned} & \int_{\mathbb{R}^d} f(t_0, x) \bar{u}(x) \psi(x) dx - \int_{\mathbb{R}^d} f_0(x) S(0, t_0) \bar{u}(x) \psi(x) dx \\ &= \int_0^{t_0} \int_{\mathbb{R}^d} f(s, x) u(s, x) b(s, x) \cdot \nabla \psi(x) dx ds. \end{aligned}$$

By linearity, the same is true for all increasing \bar{u} as well, and therefore, by a density argument, for all $\bar{u} \in L_{\text{loc}}^{p'}(\mathbb{R}^d)$. In particular, we take $\bar{u} \in C_c(\mathbb{R}^d)$, in which case $S(\cdot, t_0) \bar{u}$ is supported in $[0, T] \times B_R$ for some $R > 0$, by the finite speed of propagation property. To conclude, we may then take $\psi \in C_c^1(\mathbb{R}^d)$ such that $\psi \equiv 1$ in B_R , and therefore $\nabla \psi = 0$ in B_R . ■

Corollary 4.1. *Assume that f and $|f|$ are both distributional solutions of (4.2). Then f is the unique duality solution of (4.2); that is, $f(t, \cdot) = S(t, 0) f(0, \cdot)$.*

Corollary 4.1 gives a sufficient criterion for a distributional solution to be the correct duality solution. However, we do not know whether this condition is necessary. In other words, it is an open question whether $|S^*(t, s) \bar{f}| = S^*(t, s) |\bar{f}|$ for any $\bar{f} \in L_{\text{loc}}^p$. This renormalization property is equivalent to a kind of injectivity for the forward flow $\phi_{t,s}$, which we describe with the next result.

Proposition 4.4. *Let $0 \leq s \leq t \leq T$, $1 \leq p \leq \infty$, and $\bar{f} \in L_{\text{loc}}^p(\mathbb{R}^d)$. Then the following statements are equivalent:*

- (a) $|S^*(t, s) \bar{f}| = S^*(t, s) |\bar{f}|$.
- (b) If $A_+ := \{\bar{f} > 0\}$ and $A_- := \{\bar{f} < 0\}$, then

$$\int_{\mathbb{R}^d} \phi_{t,s}^\# \mathbf{1}_{A_+}(x) \phi_{t,s}^\# \mathbf{1}_{A_-}(x) dx = 0.$$

Proof. Let $\rho \in L_{\text{loc}}^p(\mathbb{R}^d)$ with $\rho \geq 0$, and set $A := \{x : \rho(x) > 0\}$. We first claim that

$$|\{x : S(t, s) \rho(x) > 0\} \triangle \{x : S(t, s) \mathbf{1}_A(x) > 0\}| = 0. \quad (4.14)$$

Let $B \subset \mathbb{R}^d$ be measurable. Then by Theorem 4.5,

$$\int_B S^*(t, s) \rho(x) dx = \int \rho(x) \mathbf{1}_{\{\phi_{t,s}(x) \in B\}} dx$$

and

$$\int_B S^*(t, s) \mathbf{1}_A(x) dx = |\{x \in A : \phi_{t,s}(x) \in B\}|.$$

It follows that $(S^*(t, x) \rho) \mathbf{1}_B = 0$ a.e. if and only if $(S^*(t, x) \mathbf{1}_A) \mathbf{1}_B = 0$ a.e., whence (4.14).

It now follows that (b) is equivalent to

$$|\{S^*(t, s) \bar{f}_+ \neq 0\} \cap \{S^*(t, x) \bar{f}_- \neq 0\}| = 0.$$

By linearity,

$$S^*(t, s)\bar{f} = S^*(t, s)\bar{f}_+ - S^*(t, s)\bar{f}_-,$$

and so this is equivalent to

$$(S^*(t, s)\bar{f})_{\pm} = S^*(t, s)\bar{f}_{\pm},$$

and thus (a). ■

Either of the two renormalization properties would follow from the strong convergence of regularizations in Theorem 4.5. However, we do not know at this time whether the strong convergence actually holds; see Remark 4.6.

The characterization in part (b) of Proposition 4.4 is a reformulation of the renormalization property in terms of the injectivity of the flow. For instance, we have the following:

Lemma 4.1. *Assume that b satisfies (3.2), let $0 \leq s \leq t \leq T$ and $\bar{f} \in L^p_{\text{loc}}$, and suppose that*

$$|\{x \in \mathbb{R}^d : \phi_{t,s}^{-1}(\{x\}) \cap \{\bar{f} > 0\} \neq \emptyset \text{ and } \phi_{t,s}^{-1}(\{x\}) \cap \{\bar{f} < 0\} \neq \emptyset\}| = 0.$$

Then the renormalization property in Proposition 4.4 is satisfied.

Proof. It suffices to establish part (b) in Proposition 4.4. Let $A_{\pm} = \{\pm \bar{f} > 0\}$. Then, for any measurable $B \subset \mathbb{R}^d$,

$$\int_B \phi_{t,s}^{\#} \mathbf{1}_{A_{\pm}}(x) dx = |A_{\pm} \cap \phi_{t,s}^{-1}(B)|.$$

Let B be any finite-measure set contained in

$$B_+ := \{y \in \mathbb{R}^d : \phi_{t,s}^{-1}(\{y\}) \subset A_+^c\}.$$

Then

$$\int_B \phi_{t,s}^{\#} \mathbf{1}_{A_+} dx = |A_+ \cap \phi_{t,s}^{-1}(B)| = 0.$$

It follows that $\phi_{t,s}^{\#} \mathbf{1}_{A_+} = 0$ a.e. in B_+ . Similarly, $\phi_{t,s}^{\#} \mathbf{1}_{A_-} = 0$ a.e. in

$$B_- := \{y \in \mathbb{R}^d : \phi_{t,s}^{-1}(\{y\}) \subset A_-^c\}.$$

We conclude that $\phi_{t,s}^{\#} \mathbf{1}_{A_+}(y)$ and $\phi_{t,s}^{\#} \mathbf{1}_{A_-}(y)$ are both positive only if $\phi_{t,s}^{-1}(\{y\})$ intersects both A_+ and A_- . In view of the assumption of the lemma, the set of such y has measure 0. This establishes property (b) of Proposition 4.4. ■

Remark 4.7. Observe that, when $d = 1$, $\phi_{t,s}$ is actually injective for any $0 \leq s \leq t \leq T$. For the drifts introduced in Examples 3.1 or 3.2, the corresponding flow has the property that $\phi_{t,s}^{-1}(\{y\})$ is at most a singleton for any $0 \leq s \leq t \leq T$ and a.e. $y \in \mathbb{R}^d$.

In view of the regular Lagrangian property, a kind of injectivity of the flow can be seen for particular ordered sets. Suppose that $x, y \in \mathbb{R}^d$, $x \leq y$, and $\phi_{t,s}(x) = \phi_{t,s}(y)$. Then it cannot be true that $x_i < y_i$ for all $i = 1, 2, \dots, d$. If this were the case, then $\phi_{t,s}$ would be constant on the cube $[x, y]$, which violates the regular Lagrange property. The following then follows from Lemma 4.1:

Proposition 4.5. *Assume that $\bar{f} \in L^p_{\text{loc}}(\mathbb{R}^d)$ and, for a.e. $x, y \in \mathbb{R}^d$ such that $\bar{f}(x) > 0$ and $\bar{f}(y) < 0$, either $x < y$ or $y < x$. Then renormalization is satisfied for $S^*(t, s)\bar{f}$ for all $0 \leq s \leq t \leq T$.*

The condition on \bar{f} in Proposition 4.5 is satisfied if there exist cubes $(Q_n)_{n \in \mathbb{Z}}$ such that $Q_n < Q_{n+1}$,⁴ $\{\bar{f} > 0\} \subset \bigcup_{n \in \mathbb{Z}} Q_{2n}$, and $\{\bar{f} < 0\} \subset \bigcup_{n \in \mathbb{Z}} Q_{2n+1}$.

4.3. Some remarks on “time-reversed” equations

As discussed in Section 3.4, for velocity fields b satisfying (3.2), there is no satisfactory notion of reverse flow for the ODE (3.1). Nevertheless, we can indirectly make sense of the backward Jacobian $\det(D_x \phi_{0,t}(x))$, which, formally, should be the solution of (4.2) with $f_0 = 1$, that is,

$$J(t, \cdot) := S^*(t, 0)1. \quad (4.15)$$

Proposition 4.6. *Let J be defined by (4.15):*

- (a) *For all $p \in [1, \infty)$, $J \in L^\infty \cap C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$.*
- (b) *If $(b^\varepsilon)_{\varepsilon > 0}$ are smooth, satisfy (3.2) uniformly in ε , and converge a.e. to b as $\varepsilon \rightarrow 0$, and if $\phi_{0,t}^\varepsilon$ is the solution of*

$$\partial_t \phi_{0,t}^\varepsilon(x) = -b^\varepsilon(t, \phi_{0,t}^\varepsilon(x)) \quad \text{in } [0, T], \quad \phi_{0,0}^\varepsilon(x) = x,$$

then, as $\varepsilon \rightarrow 0$, $\det(D_x \phi_{0,t}^\varepsilon)$ converges weakly in $C([0, T], L^p_{\text{loc}}(\mathbb{R}^d))$ to J .

- (c) *For all $f_0 \in L^\infty$ and $(t, x) \in [0, T] \times \mathbb{R}^d$,*

$$|S^*(t, 0)f_0(x)| \leq \|f\|_\infty J(t, x).$$

Proof. Items (a) and (b) follow immediately from Theorem 4.2. To prove (c), let b^ε and $\phi_{0,t}^\varepsilon$ be as in the statement, let f^ε be the solution of (4.2) with drift b^ε and initial condition $f_0^\varepsilon = f * \rho_\varepsilon$, where ρ_ε is a standard mollifier. Then, for $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$|f^\varepsilon(t, x)| = |f_0^\varepsilon(\phi_{0,t}^\varepsilon(x))| \det(D_x \phi_{0,t}^\varepsilon(x)).$$

The statement follows upon sending $\varepsilon \rightarrow 0$ and appealing to the weak convergence of f^ε and $\det(D_x \phi_{0,t}^\varepsilon)$. ■

⁴In other words, for all $x \in Q_n$ and $y \in Q_{n+1}$, we have $x < y$.

Continuing the formal discussion from above, note that, if f_0 and b are smooth, then $v(t, x) = f_0(\phi_{0,t}(x))$ solves the initial value problem for the transport equation

$$\partial_t v + b(t, x) \cdot \nabla v = 0 \quad \text{in } [0, T] \times \mathbb{R}^d, \quad v(0, \cdot) = f_0. \quad (4.16)$$

The time direction for (4.16) is forward, in contrast to (4.1), where it is backward. We therefore cannot appeal to the theory for that equation. Nevertheless, if $f_0 \in L^\infty$ and b satisfies (3.2), then a candidate for the solution of (4.16) is

$$v(t, x) := \frac{J(t, x)}{S^*(t, 0)f_0(x)}, \quad (t, x) \in [0, T] \times \mathbb{R}^d, \quad (4.17)$$

which, by Proposition 4.6(c), is a well-defined bounded function. Note, however, that studying the stability properties of formula (4.17) is complicated by the fact that J and $S^*(t, 0)f_0$ are stable only under weak convergence in $C([0, T], L^p_{\text{loc}})$.

To complement (4.16), we also consider the terminal value problem for the continuity equation:

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

The formula in this case should be

$$g(t, x) = g_T(\phi_{T,t}(x)) \det(D_x \phi_{T,t}(x)). \quad (4.18)$$

In fact, both terms in the product have meaning: $u(t, x) := g_T(\phi_{T,t}(x))$ is the solution of (4.1) with terminal value g_T , and $\det(D_x \phi_{T,t}(x))$ is well defined almost everywhere by Lemma 2.3 and the fact that $\phi_{T,t}$ is increasing. Furthermore, by regularizing b and taking weak distributional limits, it turns out that $\det(D_x \phi_{T,t}(x))$ is a measure bounded from below. However, u is not continuous in general, and so it is not possible to make sense of the product in (4.18). This is exactly what leads to multiple measure-valued solutions of the equation in general; see the discussion of Example 3.2 above.

Remark 4.8. The ideas described here to relate the time-reversed equations to those considered in the other sections of this paper are similar to those of Bouchut, James, and Mancini [13], and should especially be compared to their concept of reversible solutions, a theory that was expanded upon by the authors in [48]. In particular, for b satisfying a one-sided Lipschitz bound, a satisfactory solution theory can be built for both the forward and backward equations, and the two theories can be related to each other in view of the (one-sided) invertibility of the flow. This approach is limited in the current context precisely because the backward flow cannot be uniquely defined in an appropriate way.

4.4. Second-order equations

We finish this section by briefly demonstrating how the first-order results can be extended to the second-order equations

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

and

$$\partial_t f + \operatorname{div}(b(t, x)f) - \nabla^2 \cdot (a(t, x)f) = 0 \text{ in } (0, T) \times \mathbb{R}^d \quad \text{and} \quad f(0, \cdot) = f_0, \quad (4.20)$$

where b satisfies (3.2) and (3.14), and $a: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{S}^d$ is given by

$$a(t, x) = \frac{1}{2} \sigma(t, x) \sigma(t, x)^\tau$$

and σ is a matrix-valued function satisfying (3.13). Observe that this means that

$$a_{ij}(t, x) \text{ depends only on } (t, x_i, x_j) \in [0, T] \times \mathbb{R}^2 \quad \text{for all } i, j = 1, 2, \dots, d. \quad (4.21)$$

Theorem 4.7. *For all $u_T \in L^p(\mathbb{R}^d)$, $1 \leq p < \infty$, there exists a unique $u \in C([0, T], L^p)$ with the following properties:*

(a) *There exists a modulus $\omega: [0, T] \rightarrow \mathbb{R}_+$ such that*

$$\max_{t \in [0, T]} \|u(t, \cdot)\|_{L^p} \leq e^{\omega(T-t)} \|u_T\|_{L^p}.$$

(b) *If $(b^\varepsilon)_{\varepsilon>0}$ is any family of smooth functions satisfying (3.2), (3.14) uniformly in ε , converging a.e. to b as $\varepsilon \rightarrow 0$, and $(u_T^\varepsilon)_{\varepsilon>0}$ is a family of smooth functions converging in L^p to u_T , then, as $\varepsilon \rightarrow 0$, the unique solution u^ε of (4.19) converges in $C([0, T], L^p)$ to u . The same is true for vanishing viscosity limits.*

(c) *For any $(t, x) \in [0, T] \times \mathbb{R}^d$,*

$$u(t, x) = \mathbb{E}[\Phi_{T,t}(x)], \quad (4.22)$$

where Φ denotes the stochastic flow from Theorem 3.5.

Proof. The argument follows almost exactly as in the first-order case (Theorem 4.1), using the stability and uniqueness results in Theorem 3.5 for the SDE (3.12) (recall that we are assuming the bounded-oscillation condition (3.14) in addition to (3.2)). Upon regularizing the velocity field b , the formal a priori L^p estimate in part (a) can be made rigorous, which, in particular, gives boundedness of the solution operator on L^p for all $p \in [1, \infty)$, uniformly in $\varepsilon > 0$, so that the initial datum u_0 can always be assumed to belong to $C_c(\mathbb{R}^d)$ without loss of generality. The existence and uniqueness of the strong limit and its identification with the formula in part (c) are then a consequence of Theorem 3.5. ■

Remark 4.9. If u_T is increasing (decreasing), then the same is true for $u(t, \cdot)$ for all $t \in [0, T]$, which, again, can be checked with the representation formula (4.22), or by differentiating the equation and using (4.21). For now, we do not discuss the question of characterizing solutions in a PDE sense.

Remark 4.10. Similar results can be obtained for the equation with lower-order terms

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

Theorem 4.8. *For every $f_0 \in L^p$, $1 \leq p < \infty$, there exists a distributional solution of (4.20) with the following properties:*

- (a) *f is obtained uniquely from weak limits in $C([0, T], L^p)$ upon replacing b with a regularization, or from vanishing viscosity limits, and the resulting solution operator is bounded on L^p , with bound depending only on the assumptions for b , for all $p \in [1, \infty)$.*
- (b) *If $p \in [1, \infty)$, $t \in [0, T]$, and $\bar{u} \in L^{p'}$, and if u is the solution identified in Theorem 4.7 of (4.19) in $[0, t] \times \mathbb{R}^d$ with terminal value $u(t, \cdot) = \bar{u}$, then*

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

- (c) *For all $t \in [0, T]$,*

$$f(t, \cdot) = \mathbb{E}(\Phi_{t,0})_{\#} f_0,$$

where $\Phi_{t,0}$ is the stochastic flow from Theorem 3.5. Thus, if f_0 is a probability density and X_0 is a random variable independent of the Wiener process with density function f_0 , it follows that $f(t, \cdot)$ is the probability density function of $\Phi_{t,0}(X_0)$ (which is absolutely continuous with respect to Lebesgue measure in view of the regular Lagrange property).

- (d) *If $f_0 \geq 0$ and $g \geq 0$ is a nonnegative distributional solution of (4.20), then $f = g$.*

Proof. The proofs of parts (a)–(c) proceed similarly to the proof of Theorem 4.5, by first regularizing b , proving uniform L^p -estimates, and passing to the limit, exploiting the uniqueness results for (3.12). The uniqueness of nonnegative solutions in part (d) now follows from the Ambrosio–Figalli–Trevisan superposition principle; see for instance [31, 58]. ■

5. Nonlinear transport systems

We now turn to the study of the nonlinear transport systems discussed in the introduction, that is,

$$\partial_t u + f(t, x, u) \cdot \nabla_x u + g(t, x, u) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = u_T, \quad (5.1)$$

where, for some integer $m \geq 1$, $u: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m$, $f: [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^d$, and $g: [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^m$. We also consider the associated forward-backward system of characteristics: for $(t, x) \in [0, T] \times \mathbb{R}^d$ and $s \in [t, T]$, this is

$$\begin{cases} -\partial_s U_{s,t}(x) = g(s, X_{s,t}(x), U_{s,t}(x)), & U_{T,t}(x) = u_T(X_{T,t}(x)), \\ \partial_s X_{s,t}(x) = f(s, X_{s,t}(x), U_{s,t}(x)), & X_{t,t}(x) = x. \end{cases} \quad (5.2)$$

5.1. Weak solutions

We will introduce assumptions on f , g , and u_T , that, freezing u , make equation (5.1) satisfy the assumptions for the nonconservative linear equations studied in the previous section. This then leads to a natural notion of weak solution via a fixed-point operator.

Assume

$$\left\{ \begin{array}{l} (f, g): [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^d \times \mathbb{R}^m, \\ f(t, x, \cdot), g(t, x, \cdot) \text{ are continuous for } (t, x) \in [0, T] \times \mathbb{R}^d, \\ \int_0^T \sup_{(x, u) \in \mathbb{R}^d \times \mathbb{R}^m} \left(\frac{|f(t, x, u)|}{1 + |x|} + \frac{|g(t, x, u)|}{1 + |u|} \right) dt < \infty, \\ \text{and, for some } C_0 \in L^1_+([0, T]), \text{ a.e. } t \in [0, T], \\ \text{and all } i, j \in \{1, 2, \dots, d\} \text{ and } k, \ell \in \{1, 2, \dots, m\}, \\ \partial_{x_i} f^j(t, \cdot, \cdot) \geq -C_0(t) \delta_{ij}, \quad \partial_{u_k} g^\ell(t, \cdot, \cdot) \geq -C_0(t) \delta_{k\ell}, \text{ and} \\ \partial_{x_i} g^\ell \leq 0, \text{ and } \partial_{u_k} f^j \leq 0 \text{ on } [0, T] \times \mathbb{R}^d \times \mathbb{R}^m. \end{array} \right. \quad (5.3)$$

Under these assumptions, any solution operator for (5.1) should preserve the decreasing property of solutions, which we show formally assuming the data are smooth.

Proposition 5.1. *Assume f , g , and u_T are smooth with uniformly bounded derivatives, f and g satisfy (5.3), and $u_T: \mathbb{R}^d \rightarrow \mathbb{R}^m$ is decreasing with respect to (2.1). If u is a smooth solution of the terminal value problem (5.1), then, for all $t \in [0, T]$, $u(t, \cdot)$ is decreasing.*

Proof. Let $i \in \{1, 2, \dots, m\}$ and $k \in \{1, 2, \dots, d\}$ be fixed, and define $v_{ik} := \partial_k u^i$. Taking the derivative in x_k of the i -component of (5.1) gives the system

$$\begin{aligned} & \partial_t v_{ik} + f(t, x, u) \cdot \nabla v_{ik} + (\partial_{x_k} f^k(t, x, u) + \partial_{u_i} g^i(t, x, u)) v_{ik} + \partial_{x_k} g^i(t, x, u) \\ &= - \sum_{j \neq k} \partial_{x_k} f^j(t, x, u) v_{ij} - \sum_{\ell \neq i} \partial_{u_\ell} g^\ell(t, x, u) v_{\ell k} - \sum_{j=1}^d \sum_{\ell=1}^m \partial_{u_\ell} f^j(t, x, u) v_{\ell k} v_{ij}. \end{aligned}$$

In view of (5.3), the system satisfied by $\{v_{ik} : i = 1, 2, \dots, m \text{ and } k = 1, 2, \dots, d\}$, after reversing time, is of the form in (2.5), and so the result follows from Lemma 2.6. ■

We now make the connection with the linear transport equation theory of the previous sections. In particular, note that if $u(t, \cdot)$ is decreasing for all $t \in [0, T]$, then $b(t, x) := f(t, x, u(t, x))$ satisfies the assumptions of (3.2), while

$$c(t) = C_0(t) \quad \text{and} \quad d(t, x) = g(t, x, u(t, x)) + C_0(t) u(t, x)$$

satisfy (4.10).

Definition 5.1. Assume f and g satisfy (5.3) and u_T is decreasing. Then a locally bounded function $u: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ that is decreasing in the \mathbb{R}^d -variable and satisfies

$$\left\{ (t, x) \mapsto \frac{u(t, x)}{1 + |x|} \right\} \in L^\infty \quad (5.4)$$

is called a solution of (5.1) if u is a solution of the linear equation (4.7) with $b(t, x) = f(t, x, u(t, x))$, $c(t) = C_0(t)$, and $d(t, x) = g(t, x, u(t, x)) + C_0(t)u(t, x)$.

Solving (5.1) in the sense of Definition 5.1 thus amounts to solving a fixed-point problem. We similarly give a weak sense to system (5.2) by using the properties of the flow from Section 3.

Proposition 5.2. Assume that $u_T \in L^\infty_{\text{loc}}$ and u is a solution of (5.1) in the sense of Definition 5.1. Let ϕ be the forward regular Lagrangian flow as in Section 3 corresponding to

$$\partial_t \phi_{t,s}(x) = f(t, \phi_{t,s}(x), u(t, \phi_{t,s}(x))), \quad t \in [s, T], \quad \phi_{s,s}(x) = x,$$

and define, for $0 \leq s \leq t \leq T$ and $x \in \mathbb{R}^d$,

$$X_{t,s}(x) = \phi_{t,s}(x) \quad \text{and} \quad U_{t,s}(x) := u(t, \phi_{t,s}(x)).$$

Then, for all $1 \leq p < \infty$, $X_{\cdot,s}, U_{\cdot,s} \in C([s, T], L^p_{\text{loc}}) \cap L^p_{\text{loc}}(\mathbb{R}^d, W^{1,1}([0, T]))$, and, for a.e. $x \in \mathbb{R}^d$, (5.2) is satisfied in the integral sense.

Proof. The regularity properties of X are seen from Theorem 3.1, as well as the fact that, for a.e. $x \in \mathbb{R}^d$,

$$X_{t,s}(x) = x + \int_s^t f(r, X_{r,s}(x), u(r, X_{r,s}(x))) dr = x + \int_s^t f(r, X_{r,s}(x), U_{r,s}(x)) dr.$$

Theorems 3.1 and 4.3 together imply that $U_{\cdot,s} \in C([s, T], L^p_{\text{loc}})$. Also, in view of Theorem 4.3 (c), u satisfies, for any $0 \leq s \leq t \leq T$,

$$\begin{aligned} U_{s,s}(x) &= u(s, x) = u(t, X_{t,s}(x)) + \int_s^t g(r, X_{r,s}(x), u(r, \phi_{r,s}(x))) dr \\ &= U_{t,s}(x) + \int_s^t g(r, X_{r,s}(x), U_{r,s}(x)) dr. \end{aligned}$$

It follows that, in the distributional sense,

$$\partial_t U_{t,s}(x) = g(t, X_{t,s}(x), U_{t,s}(x)).$$

Arguing as in the proof of Theorem 3.1, the right-hand side belongs to $L^p_{\text{loc}}(\mathbb{R}^d, L^1([0, T]))$, and we conclude as in that theorem. ■

5.2. Minimal and maximal solutions

In this section we show that assumptions (5.3) give an increasing structure to equation (5.1), which allows for the identification of unique minimal and maximal solutions.

Theorem 5.1. *Assume f and g satisfy (5.3), u_T is decreasing, and $u_T(1 + |\cdot|)^{-1} \in L^\infty$. Then there exist two decreasing solutions u^+ , u^- of (5.1) in the sense of Definition 5.1 with the following properties:*

- (a) *For all $t \in [0, T]$, $u^+(t, \cdot)$ is upper semicontinuous, $u^-(t, \cdot)$ is lower semicontinuous, and both u^+ and u^- are continuous a.e. in $[0, T] \times \mathbb{R}^d$.*
- (b) *If u is any other solution, then $u^- \leq u \leq u^+$.*

Throughout this subsection, we may assume, without loss of generality, that the term $C_0 \in L^1_+$ in (5.3) is 0. Indeed, the function

$$\tilde{u}(t, x) = \exp\left(\int_t^T C_0(s) ds\right) u\left(t, \exp\left(\int_t^T C_0(s) ds\right) x\right)$$

is decreasing in x , satisfies (5.4), and, formally, solves equation (5.1) with f and g replaced by

$$\begin{aligned} \tilde{f}(t, x, u) &= \exp\left(-\int_t^T C_0(s) ds\right) f\left(t, \exp\left(\int_t^T C_0(s) ds\right) x, \exp\left(-\int_t^T C_0(s) ds\right) u\right) \\ &\quad + C_0(t)x \end{aligned}$$

and

$$\begin{aligned} \tilde{g}(t, x, u) &= \exp\left(\int_t^T C_0(s) ds\right) g\left(t, \exp\left(\int_t^T C_0(s) ds\right) x, \exp\left(-\int_t^T C_0(s) ds\right) u\right) \\ &\quad + C_0(t)u. \end{aligned}$$

Proof of Theorem 5.1. For some $\bar{C} > 0$ to be determined, set

$$\begin{aligned} \mathcal{L} := \{u: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m : u(T, \cdot) = u_T, u(t, \cdot) \text{ is decreasing for all } t \in [0, T], \\ \text{and } |u(t, x)| \leq \bar{C}(1 + |x|) \text{ for all } (t, x) \in [0, T] \times \mathbb{R}^d\}. \end{aligned}$$

We define a map \mathcal{S} on \mathcal{L} as follows: for $u \in \mathcal{L}$, let $v := \mathcal{S}(u)$ be the solution, as in Theorem 4.3, of the linear transport equation

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

Then, in view of the solution formula (4.9) and the bounds (5.3) on f and g , there exists a sufficiently large \bar{C} , depending on u_T , such that \mathcal{S} maps \mathcal{L} into \mathcal{L} .

We now note that \mathcal{L} forms a *complete lattice* under the partial order

$$u \leq \tilde{u} \Leftrightarrow u^i(t, x) \leq \tilde{u}^i(t, x) \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d, i = 1, 2, \dots, m; \quad (5.5)$$

that is, every subset of \mathcal{L} has a greatest lower bound and least upper bound, which is a consequence of the uniformly bounded linear growth of solutions in \mathcal{L} .

Suppose now that $u, \tilde{u} \in \mathcal{L}$ satisfy $u \leq \tilde{u}$ under the order (5.5), and set

$$b(t, x) := f(t, x, u(t, x)) \quad \text{and} \quad d(t, x) := g(t, x, u(t, x)).$$

Then (5.3) with $C_0 \equiv 0$ implies that $b(t, x) \geq f(t, x, \tilde{u}(t, x))$ and $d(t, x) \leq g(t, x, \tilde{u}(t, x))$, and so, in particular, $v := \mathcal{S}(u)$ and $\tilde{v} := \mathcal{S}(\tilde{u})$ are respectively a sub- and a supersolution of the linear equation (4.7) with $c(t) \equiv 0$. It follows from Theorem 4.4 that $v \leq \tilde{v}$, and therefore \mathcal{S} is increasing on the complete lattice \mathcal{L} with respect to the partial order (5.5). The existence of unique maximal and minimal solutions is now a consequence of the Tarski lattice-theoretical fixed-point theorem [57].

The continuity a.e. in $[0, T] \times \mathbb{R}^d$ of u^+ and u^- now follows from Theorem 4.4. Observe now that any version of the maximal solution u^+ is a solution in the sense of Definition 5.1. Because $u^+(t, \cdot)$ is decreasing, it is continuous a.e., and its maximal version is upper semicontinuous, which is therefore also the unique maximal solution in the *everywhere*-pointwise sense. A similar argument shows that the minimal everywhere-pointwise solution u^- is lower semicontinuous in the spatial variable, and we conclude. ■

The fixed-point theorem of [57] used above further characterizes u^+ and u^- as

$$u^+ = \sup\{u \in \mathcal{L} : \mathcal{S}(u) \geq u\} \quad \text{and} \quad u^- = \inf\{u \in \mathcal{L} : \mathcal{S}(u) \leq u\},$$

where the sup and inf are understood with respect to the partial order (5.5). We alternatively characterize the maximal and minimal solutions in terms of appropriately defined sub- and supersolutions of equation (5.1).

Recall that ρ^ε and ρ_ε are the one-sided mollifying functions from Section 2.3.

Definition 5.2. Assume f and g satisfy (5.3) with $C_0 \equiv 0$ and u_T is decreasing. A function $u: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ that is decreasing in \mathbb{R}^d is called a subsolution (resp. supersolution) of (5.1) if $u(T, \cdot) \leq u_T$ (resp. $u(T, \cdot) \geq u_T$) and $u^\varepsilon := u * \rho^\varepsilon$ (resp. $u_\varepsilon := u * \rho_\varepsilon$) satisfies

$$\begin{aligned} & -\partial_t u^\varepsilon - f(t, x, u(t, x)) \cdot \nabla u^\varepsilon - g(t, x, u(t, x)) \leq 0 \\ & (\text{resp. } -\partial_t u_\varepsilon - f(t, x, u(t, x)) \cdot \nabla u_\varepsilon - g(t, x, u(t, x)) \geq 0) \quad \text{a.e. in } [0, T] \times \mathbb{R}^d. \end{aligned}$$

In other words, sub- and supersolutions in the sense of (5.2) are sub- and supersolutions of the corresponding linear equations identified in Definition 5.1. It follows that a function which is both a sub- and a supersolution is in fact a solution.

Lemma 5.1. *Under the conditions of Definition 5.2, assume that u and v are two subsolutions (resp. supersolutions) of (5.1). Then $u \vee v$ (resp. $u \wedge v$) is also a subsolution (resp. supersolution).*

Proof. We prove only the subsolution statement, as the other one is analogously proved. It is clear that $\max\{u(T, \cdot), v(T, \cdot)\} \leq u_T$ and $u \vee v$ is decreasing in \mathbb{R}^d .

Set $u^\varepsilon = u * \rho^\varepsilon$ and $v^\varepsilon = v * \rho^\varepsilon$. Let $\Psi: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ be smooth, increasing, and satisfy $\Psi(u, v) \geq u \vee v$ for all $u, v \in \mathbb{R}^m$. Note that, using (5.3) with $C_0 \equiv 0$,

$$f(t, x, u(t, x)) \wedge f(t, x, v(t, x)) \geq f(t, x, u(t, x) \vee v(t, x))$$

and

$$g(t, x, u(t, x)) \vee g(t, x, v(t, x)) \leq g(t, x, u(t, x) \vee v(t, x)),$$

and so a simple application of the chain rule gives

$$\partial_t \Psi(u^\varepsilon, v^\varepsilon) + f(t, x, u \vee v) \cdot \nabla \Psi(u^\varepsilon, v^\varepsilon) + g(t, x, u \vee v) \geq 0.$$

In particular, for any smooth positive test function $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$,

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} (\Psi(u^\varepsilon, v^\varepsilon)(t, x) \partial_t \phi(t, x) - g(t, x, u \vee v) \phi(t, x)) dx dt \\ + \langle \nabla(f(\cdot, \cdot, u \vee v) \phi), \Psi(u^\varepsilon, v^\varepsilon) \rangle \leq 0, \end{aligned}$$

where the last term is understood as the pairing between locally finite measures and continuous functions, because $f(t, x, (u \vee v)(t, x))$ is BV in x . We may then approximate $\max(\cdot, \cdot)$ with such functions Ψ and determine that, in the distributional sense,

$$\partial_t(u^\varepsilon \vee v^\varepsilon) + f(t, x, u \vee v) \cdot \nabla(u^\varepsilon \vee v^\varepsilon) + g(t, x, u \vee v) \geq 0.$$

For $\delta > 0$, we convolve both sides of the above inequality with the one-sided mollifier ρ^δ and obtain, in view of (5.3),

$$\partial_t(u^\varepsilon \vee v^\varepsilon) * \rho^\delta + f(t, x, u \vee v) \cdot \nabla[(u^\varepsilon \vee v^\varepsilon) * \rho^\delta] + g(t, x, u \vee v) \geq 0.$$

For fixed δ , as $\varepsilon \rightarrow 0$, $(u^\varepsilon \vee v^\varepsilon) * \rho^\delta \rightarrow (u \vee v) * \rho^\delta$ locally uniformly, and we may therefore send $\varepsilon \rightarrow 0$ in the above inequality, again using $f \in \text{BV}$, to obtain the desired subsolution inequality for $(u \vee v) * \rho^\delta$. ■

Lemma 5.2. Assume f and g satisfy (5.3) with $C_0 \equiv 0$, u_T is decreasing, and u is a subsolution (supersolution) of (5.1) in the sense of Definition 5.2. Then there exists a subsolution (supersolution) \tilde{u} such that $u \leq \tilde{u}$ ($u \geq \tilde{u}$) and \tilde{u} is continuous a.e. in $[0, T] \times \mathbb{R}^d$.

Proof. For such a subsolution u , let \tilde{u} be the solution of the linear transport equation

$$\partial_t \tilde{u} + f(t, x, u(t, x)) \cdot \nabla \tilde{u} + g(t, x, u(t, x)) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \quad \tilde{u}(T, \cdot) = u_T.$$

Then, by Theorem 4.4, $u \leq \tilde{u}$, and \tilde{u} is continuous a.e. in $[0, T] \times \mathbb{R}^d$. By (5.3),

$$f(t, x, u(t, x)) \geq f(t, x, \tilde{u}(t, x)) \quad \text{and} \quad g(t, x, u(t, x)) \leq g(t, x, \tilde{u}(t, x)),$$

and it follows that \tilde{u} is a subsolution of (5.1). The argument for supersolutions is analogous. ■

Proposition 5.3. *Let f and g satisfy (5.3) with $C_0 \equiv 0$, and assume u_T has linear growth and is decreasing. Let u^+ and u^- be the maximal and minimal solution from Theorem 5.1. Then, in the sense of Definition 5.2, u^+ is the pointwise maximum of all subsolutions, and u^- is the pointwise minimum of all supersolutions.*

Remark 5.1. In view of Lemma 5.2, the maximum/minimum in Proposition 5.3 may be restricted to sub/supersolutions that are continuous a.e. in $[0, T] \times \mathbb{R}^d$.

Proof of Proposition 5.3. We prove only the statement for u^+ , since the proof is analogous for u^- . Let

$$\tilde{u}^+(t, x) := \sup\{u(t, x) : u \text{ is a subsolution}\}.$$

Because u^+ is itself a subsolution, we clearly have $u^+ \leq \tilde{u}^+$. Suppose now that u is a subsolution of (5.1), and let v be the solution of

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

It then holds that u and v are respectively a subsolution and solution of the linear equation (4.7) with $b(t, x) = f(t, x, u(t, x))$, $c(t, x) = 0$, and $d(t, x, \cdot) = g(t, x, u(t, x))$. Then Theorem 4.4 implies $u \leq v$. Note also that v belongs to the lattice \mathcal{L} from the proof of Theorem 5.1, because $v(T, \cdot) = u_T$. If \mathcal{S} is the fixed-point map from the proof of that theorem, we then set $w = \mathcal{S}(v)$, which, by a similar argument, satisfies $w \in \mathcal{L}$ and $\mathcal{S}(v) = w \geq v$. By the characterization of u^+ by the Tarski fixed-point theorem, we must have $u \leq u^+$, and therefore $\tilde{u}^+ \leq u^+$. ■

5.3. Continuous solutions

We now investigate when the maximal and minimal solutions u^+ and u^- identified in the previous subsection coincide. In this subsection we prove a comparison principle for *continuous* sub- and supersolutions, so that, in particular, $u^+ = u^-$ if both are continuous. As we will see by example in the forthcoming subsections, this can fail in general if the assumption of continuity is dropped.

We will first introduce a different notion of solutions, which will turn out to be equivalent when the solutions are continuous, using the theory of viscosity solutions. The definition of solution we give here, which allows for discontinuities, is similar to the one introduced by Ishii [39] in the context of Perron's method.

Let us assume throughout this subsection, in addition to (5.3), that

$$\begin{aligned} &f \text{ and } g \text{ are uniformly Lipschitz continuous and bounded,} \\ &\text{and } C_0 \geq 0 \text{ is constant in } t \in [0, T]. \end{aligned} \tag{5.6}$$

Definition 5.3. A function $u: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ is a viscosity subsolution (supersolution) of (5.1) if $u(t, \cdot)$ is decreasing for all $t \in [0, T]$, and, whenever $\phi \in C^1([0, T] \times \mathbb{R}^d, \mathbb{R})$,

$i = 1, 2, \dots, m$, and $u^{i,*}(t, x) - \phi(t, x)$ (resp. $u_*^i(t, x) - \phi(t, x)$) attains a local maximum (resp. minimum) at (t_0, x_0) ,

$$\begin{aligned} & -\phi_t(t_0, x_0) - f(t_0, x_0, u^*(t_0, x_0)) \cdot \nabla \phi(t_0, x_0) - g(t_0, x_0, u^*(t_0, x_0)) \leq 0 \\ & \text{(resp. } -\phi_t(t_0, x_0) - f(t_0, x_0, u_*(t_0, x_0)) \cdot \nabla \phi(t_0, x_0) - g(t_0, x_0, u_*(t_0, x_0)) \geq 0). \end{aligned}$$

A viscosity solution is both a sub- and a supersolution.

Remark 5.2. Due to the evolutionary nature of the equation (in reversed time), in Definition 5.3, it suffices to interpret local extrema (t_0, x_0) over half-open neighborhoods of the form $[t_0, t_0 + h) \times B_r(x_0)$ for some $h, r > 0$; see for instance the argument in [24, Theorem 8.2].

Lemma 5.3. Assume (5.3) and (5.6). If u is a subsolution (supersolution) in the sense of Definition 5.1 and is almost-everywhere continuous in $[0, T] \times \mathbb{R}^d$, then u is a viscosity subsolution (supersolution) in the sense of Definition 5.3.

Remark 5.3. In view of Lemma 5.2, no generality is lost in considering sub- or supersolutions that are continuous a.e. in $[0, T] \times \mathbb{R}^d$.

Proof of Lemma 5.3. We prove only the subsolution property. Note that, because u is decreasing, we have $u(t, \cdot)^* = u(t, \cdot)$ a.e. Then $u^\varepsilon = u * \rho^\varepsilon$ is a (classical) subsolution of

$$-\partial_t u^\varepsilon - f(t, x, u^*(t, x)) \cdot \nabla u^\varepsilon - g(t, x, u^*(t, x)) \leq 0,$$

and, as $\varepsilon \rightarrow 0$, $u^\varepsilon \searrow u^*$.

Assume that $u^*(t, x) - \phi(t, x)$ attains a maximum at (t_0, x_0) , which we may assume to be strict by adding a quadratic penalization to ϕ . For $\varepsilon > 0$, let $(t_\varepsilon, x_\varepsilon)$ denote the maximum point of $u^\varepsilon(t, x) - \phi(t, x)$ on $[0, T] \times \overline{B_1(x_0)}$, and let (s, y) be a limit point of $(t_\varepsilon, x_\varepsilon)_{\varepsilon > 0}$ along some subsequence $(\varepsilon_n)_{n \in \mathbb{N}}$. By Lemma 2.5 (c),

$$u^\varepsilon(t_\varepsilon, x_\varepsilon) \leq u(t_\varepsilon, x_\varepsilon - 2\varepsilon \mathbf{1}) + \varepsilon \mathbf{1},$$

and so

$$\limsup_{n \rightarrow \infty} u^{\varepsilon_n}(t_{\varepsilon_n}, x_{\varepsilon_n}) \leq u^*(s, y).$$

Now let $(s_n, y_n)_{n \in \mathbb{N}} \subset [0, T] \times \overline{B_1(x_0)}$ be such that

$$(s_n, y_n) \xrightarrow{n \rightarrow \infty} (t_0, x_0) \quad \text{and} \quad u^{\varepsilon_n}(s_n, y_n) \xrightarrow{n \rightarrow \infty} u^*(t_0, x_0).$$

Taking $n \rightarrow \infty$ in the inequality $u^{\varepsilon_n}(s_n, y_n) - \phi(s_n, y_n) \leq u^{\varepsilon_n}(t_{\varepsilon_n}, x_{\varepsilon_n}) - \phi(t_{\varepsilon_n}, x_{\varepsilon_n})$ yields $u^*(t_0, y_0) - \phi(t_0, y_0) \leq u^*(s, y) - \phi(s, y)$. This implies that $(s, y) = (t_0, x_0)$, the limit holds along the full family $\varepsilon \rightarrow 0$, and $\lim_{\varepsilon \rightarrow 0} u^\varepsilon(t_\varepsilon, x_\varepsilon) = u^*(t_0, x_0)$.

By standard maximum principle considerations, and the facts that $-f$ and g are increasing in u and $\nabla u^\varepsilon \leq 0$,

$$-\partial_t \phi(t_\varepsilon, x_\varepsilon) - f(t_\varepsilon, x_\varepsilon, u^\varepsilon(t_\varepsilon, x_\varepsilon)) \cdot \nabla \phi(t_\varepsilon, x_\varepsilon) - g(t_\varepsilon, x_\varepsilon, u^\varepsilon(t_\varepsilon, x_\varepsilon)) \leq 0.$$

Sending $\varepsilon \rightarrow 0$ gives the desired subsolution inequality. ■

Theorem 5.2. Assume that f and g satisfy (5.3) and (5.6), and let u and v be respectively a bounded sub- and supersolution of (5.1) in the sense of Definition 5.3 such that either u is continuous and v is lower semicontinuous, or u is upper semicontinuous and v is continuous. If $u(T, \cdot) \leq v(T, \cdot)$, then, for all $t \in [0, T]$, $u(t, \cdot) \leq v(t, \cdot)$.

Proof. The proofs of both statements are almost identical, so we prove only the statement when u is continuous and v is lower semicontinuous.

We first prove the result under the additional assumption that, for some $\delta > 0$, u is a subsolution of

$$-\partial_t u - [f(t, x, u) + \delta \mathbf{1}] \cdot \nabla u - g(t, x, u) \leq -\delta \mathbf{1} \quad (5.7)$$

and $u(T, \cdot) - v(T, \cdot) \leq -\delta \mathbf{1}$. Standard arguments from the theory of viscosity solutions then imply that $w(t, x, y) := u(t, x) - v(t, y)$ is a subsolution of

$$-\partial_t w - [f(t, x, u) + \delta \mathbf{1}] \cdot \nabla_x w - f(t, y, v) \cdot \nabla_y w - g(t, x, u) + g(t, y, v) \leq -\delta \mathbf{1}. \quad (5.8)$$

Fix $\lambda \geq 1$ and $\beta > 0$, set

$$[0, T] \ni t \mapsto \phi(t) := \max_{x, y \in \mathbb{R}^d} \max_{i=1,2,\dots,m} \left(u^i(t, x) - v^i(t, y) - \frac{\lambda}{2} |x - y|^2 - \frac{\beta}{2} (|x|^2 + |y|^2) \right),$$

and define

$$\hat{t} := \sup \{ t \in (-\infty, T] : \phi(t) > 0 \}.$$

As $\lambda \rightarrow \infty$,

$$\phi(T) = \max_{x, y, i} \left(u^i(T, x) - v^i(T, y) - \frac{\lambda}{2} |x - y|^2 - \frac{\beta}{2} (|x|^2 + |y|^2) \right) \leq -\delta + o(1),$$

and so, if λ is sufficiently large, depending on δ , then $\phi(T) < 0$, and therefore $\hat{t} < T$.

We next claim that $\hat{t} < 0$ for all sufficiently large λ and small β , which implies $\phi(t) \leq 0$ for all $t \in [0, T]$, and, hence, gives the result. If this were not true and $0 \leq \hat{t} < T$, then choose i, \hat{x}, \hat{y} such that

$$0 = \phi(\hat{t}) = u^i(\hat{t}, \hat{x}) - v^i(\hat{t}, \hat{y}) - \frac{\lambda}{2} |\hat{x} - \hat{y}|^2 - \frac{\beta}{2} (|\hat{x}|^2 + |\hat{y}|^2).$$

We then have

$$u^j(\hat{t}, \hat{x}) - v^j(\hat{t}, \hat{y}) \leq u^i(\hat{t}, \hat{x}) - v^i(\hat{t}, \hat{y}) \quad \text{for all } j = 1, 2, \dots, m. \quad (5.9)$$

By definition of \hat{x} and \hat{y} , the functions

$$x \mapsto u^i(\hat{t}, x) - \frac{\lambda}{2} |x - \hat{y}|^2 - \frac{\beta}{2} |x|^2 \quad \text{and} \quad y \mapsto -v^i(\hat{t}, y) - \frac{\lambda}{2} |\hat{x} - y|^2 - \frac{\beta}{2} |y|^2$$

achieve their maxima at respectively $x = \hat{x}$ and $y = \hat{y}$, and so, from the fact that u and v are decreasing, we have

$$\lambda(\hat{x}_j - \hat{y}_j) + \beta \hat{x}_j \leq 0 \quad \text{and} \quad \lambda(\hat{x}_j - \hat{y}_j) - \beta \hat{y}_j \leq 0 \quad \text{for all } j = 1, 2, \dots, d. \quad (5.10)$$

Moreover, in view of the boundedness of u and v , arguing as in for instance [24, Lemma 3.1], there exist δ_λ and ε_β satisfying $\lim_{\lambda \rightarrow \infty} \delta_\lambda = \lim_{\beta \rightarrow 0} \varepsilon_\beta = 0$ such that

$$\lambda|\hat{x} - \hat{y}|^2 \leq \delta_\lambda \quad \text{and} \quad \beta(|\hat{x}|^2 + |\hat{y}|^2) \leq \varepsilon_\beta. \quad (5.11)$$

In particular,

$$0 \leq u^i(\hat{t}, \hat{x}) - v^i(\hat{t}, \hat{y}) \leq 2\delta_\lambda + 2\varepsilon_\beta. \quad (5.12)$$

On $[\hat{t}, T] \times \mathbb{R}^d \times \mathbb{R}^d$, the function

$$u^i(t, x) - v^i(t, y) - \frac{\lambda}{2}|x - y|^2 - \frac{\beta}{2}(|x|^2 + |y|^2)$$

achieves a maximum at $(\hat{t}, \hat{x}, \hat{y})$, and so, applying Definition 5.3 to the doubled equation (5.8) (see Remark 5.2),

$$\begin{aligned} \delta &\leq \sum_{j=1}^d (f^j(\hat{t}, \hat{x}, u(\hat{t}, \hat{x})) + \delta)(\lambda(\hat{x}_j - \hat{y}_j) + \beta\hat{x}_j) + g^i(\hat{t}, \hat{x}, u(\hat{t}, \hat{x})) \\ &\quad - \sum_{j=1}^d f^j(\hat{t}, \hat{y}, v(\hat{t}, \hat{y}))(\lambda(\hat{x}_j - \hat{y}_j) - \beta\hat{y}_j) - g^i(\hat{t}, \hat{y}, v(\hat{t}, \hat{y})). \end{aligned} \quad (5.13)$$

Because of (5.3), f is nonincreasing in the u -variable, and so the inequalities (5.9) and (5.10), together with the Lipschitz regularity of f and the inequality $\sum_{i=1}^d |x_i| \geq d^{-1/2}|x|$, imply

$$\begin{aligned} &\sum_{j=1}^d [f^j(\hat{t}, \hat{x}, u(\hat{t}, \hat{x})) + \delta](\lambda(\hat{x}_j - \hat{y}_j) + \beta\hat{x}_j) \\ &\leq \sum_{j=1}^d [f^j(\hat{t}, \hat{x}, v(\hat{t}, \hat{y})) + (u^i(\hat{t}, \hat{x}) - v^i(\hat{t}, \hat{y}))\mathbf{1} + \delta](\lambda(\hat{x}_j - \hat{y}_j) + \beta\hat{x}_j) \\ &\leq -\frac{1}{d^{1/2}}\delta\lambda|\hat{x} - \hat{y}| + \sum_{j=1}^d f^j(\hat{t}, \hat{y}, v(\hat{t}, \hat{y}))(\lambda(\hat{x}_j - \hat{y}_j) - \beta\hat{y}_j) \\ &\quad + \|\nabla_u f\|_\infty \lambda|\hat{x} - \hat{y}|(u^i(\hat{t}, \hat{x}) - v^i(\hat{t}, \hat{y})) + \|\nabla_x f\|_\infty \lambda|\hat{x} - \hat{y}|^2 \\ &\quad + (\delta + \|f\|_\infty)\beta|\hat{x}| + \|f\|_\infty \beta|\hat{y}|. \end{aligned}$$

The inequalities (5.11) and (5.12) thus give

$$\begin{aligned} &\sum_{j=1}^d [f^j(\hat{t}, \hat{x}, u(\hat{t}, \hat{x})) + \delta](\lambda(\hat{x}_j - \hat{y}_j) + \beta\hat{x}_j) \\ &\leq \sum_{j=1}^d f^j(\hat{t}, \hat{y}, v(\hat{t}, \hat{y}))(\lambda(\hat{x}_j - \hat{y}_j) - \beta\hat{y}_j) \\ &\quad + \lambda|\hat{x} - \hat{y}| \left(-\frac{\delta}{d^{1/2}} + 2\|\nabla_u f\|_\infty(\delta_\lambda + \varepsilon_\beta) \right) \\ &\quad + \|\nabla_x f\|_\infty \delta_\lambda + (\delta + 2\|f\|_\infty)\beta^{1/2}\varepsilon_\beta^{1/2}. \end{aligned}$$

Similarly, using the fact that g^i is nondecreasing in the u_j -variable for all $j \neq i$, we have

$$\begin{aligned} g^i(\hat{t}, \hat{x}, u(\hat{t}, \hat{x})) &\leq g(\hat{t}, \hat{y}, v(\hat{t}, \hat{y})) + \|\nabla_u g\|_\infty (u^i(\hat{t}, \hat{x}) - v^i(\hat{t}, \hat{y})) + \|\nabla_x g\|_\infty |\hat{x} - \hat{y}| \\ &\leq g(\hat{t}, \hat{y}, v(\hat{t}, \hat{y})) + 2\|\nabla_u g\|_\infty (\delta_\lambda + \varepsilon_\beta) + \|\nabla_x g\|_\infty \lambda^{-1/2} \delta_\lambda^{1/2}. \end{aligned}$$

Combining this with (5.13) then yields

$$\begin{aligned} \delta &\leq \lambda |\hat{x} - \hat{y}| \left(-\frac{\delta}{d^{1/2}} + 2\|\nabla_u f\|_\infty (\delta_\lambda + \varepsilon_\beta) \right) \\ &\quad + \|\nabla_x f\|_\infty \delta_\lambda + (\delta + 2\|f\|_\infty) \beta^{1/2} \varepsilon_\beta^{1/2} \\ &\quad + 2\|\nabla_u g\|_\infty (\delta_\lambda + \varepsilon_\beta) + \|\nabla_x g\|_\infty \lambda^{-1/2} \delta_\lambda^{1/2} \\ &\leq \|\nabla_x f\|_\infty \delta_\lambda + (\delta + 2\|f\|_\infty) \beta^{1/2} \varepsilon_\beta^{1/2} \\ &\quad + 2\|\nabla_u g\|_\infty (\delta_\lambda + \varepsilon_\beta) + \|\nabla_x g\|_\infty \lambda^{-1/2} \delta_\lambda^{1/2}, \end{aligned}$$

where the final inequality follows if λ is taken large enough and β is taken small enough, depending on δ , so that $2d^{1/2}\|\nabla_u f\|_\infty(\delta_\lambda + \varepsilon_\beta) \leq \delta$. Further taking λ large and β small then yields a contradiction.

We now prove the general statement. For $\delta > 0$ and $(t, x) \in [0, T] \times \mathbb{R}^d$, set

$$\tilde{u}(t, x) = u(t, x + \delta\psi_1(t)\mathbf{1}) - \delta\psi_2(t)\mathbf{1},$$

where ψ_1 and ψ_2 are two nonnegative scalar functions satisfying $\psi_1(T) = 0$ and $\psi_2(T) = 1$. Then

$$\tilde{u}(T, \cdot) = u(T, \cdot) - \delta\mathbf{1},$$

and so, using the fact that u is nonincreasing, we formally compute⁵

$$\begin{aligned} &-\partial_t \tilde{u} - [f(t, x, \tilde{u}) - \delta]\nabla \tilde{u} - g(t, x, \tilde{u}) \\ &= \partial_t u - \delta \nabla_x u \cdot \mathbf{1} \dot{\psi}_1 + \delta \dot{\psi}_2 \mathbf{1} - [f(t, x + \delta\psi_1 \mathbf{1}, u - \delta\psi_2 \mathbf{1}) + \delta]\nabla u \\ &\quad - g(x + \delta\psi_1 \mathbf{1}, u - \delta\psi_2 \mathbf{1}) + \delta \mathbf{1} \\ &\leq \delta(-\mathbf{1} \cdot \nabla u)(\dot{\psi}_1 + \|\nabla_u f\|_\infty \psi_2 + 1) + \delta(\dot{\psi}_2 + \|\nabla_u g\|_\infty \psi_2 + 1)\mathbf{1}. \end{aligned} \quad (5.14)$$

We then choose ψ_1 and ψ_2 so as to satisfy

$$\dot{\psi}_1 = -\|\nabla_u f\|_\infty \psi_2 - 1 \quad \text{and} \quad \dot{\psi}_2 = -\|\nabla_u g\|_\infty \psi_2 - 1,$$

that is,

$$\psi_1(t) = \frac{\|\nabla_u f\|_\infty (\|\nabla_u g\|_\infty + 1)}{\|\nabla_u g\|_\infty^2} (e^{\|\nabla_u g\|_\infty (T-t)} - 1) + \left(1 - \frac{\|\nabla_u f\|_\infty}{\|\nabla_u g\|_\infty}\right) (T - t)$$

⁵The computations can be made rigorous using test functions and Definition 5.3. Note that the argument for u above is always $(t, x + \delta\psi_1(t)\mathbf{1})$.

and

$$\psi_2(t) = \left(1 + \frac{1}{\|\nabla_u g\|_\infty}\right) e^{\|\nabla_u g\|_\infty(T-t)} - \frac{1}{\|\nabla_u g\|_\infty}.$$

Then (5.14) becomes exactly (5.7), and so, for $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$u(t, x + \delta\psi_1(t)) - \delta\psi_2(t) \leq v(t, x).$$

We conclude upon sending $\delta \rightarrow 0$ and appealing to the continuity of u . ■

Let us note the following corollary of Theorem 5.1, Lemma 5.3, and Theorem 5.2.

Corollary 5.1. *Under the same conditions on f and g as in Theorem 5.2, let u_T be bounded and decreasing and let u^+ and u^- be the maximal and minimal solutions identified in Theorem 5.1. Then*

$$u^+(t, x) = \inf\{u(t, x) : u \text{ is a continuous viscosity supersolution with } u(T, \cdot) \geq u_T\}$$

and

$$u^-(t, x) = \sup\{u(t, x) : u \text{ is a continuous viscosity subsolution with } u(T, \cdot) \leq u_T\}.$$

The comparison principle with continuous sub- and supersolutions also implies a conditional uniqueness and stability statement.

Theorem 5.3. *Assume f and g satisfy (5.3) and (5.6) and u_T is bounded, continuous, and decreasing. If there exists a bounded and continuous viscosity solution u of (5.1), then it is the unique viscosity solution. Moreover, if u^ε is the unique classical solution of*

$$\partial_t u^\varepsilon + f(t, x, u^\varepsilon) \cdot \nabla u^\varepsilon + g(t, x, u^\varepsilon) + \varepsilon \Delta u^\varepsilon = 0 \text{ in } (0, T) \times \mathbb{R}^d, \quad u^\varepsilon(T, \cdot) = u_T,$$

then, as $\varepsilon \rightarrow 0$, u^ε converges locally uniformly to u .

Proof. The uniqueness is a consequence of the comparison principle in Theorem 5.2, since u is both a sub- and a supersolution. The local uniform convergence as $\varepsilon \rightarrow 0$ of u^ε to the unique continuous viscosity solution u is then proved with standard stability arguments from the theory of viscosity solutions, arguing with half-relaxed limits and applying the comparison principle, Theorem 5.2. ■

5.4. A one-dimensional example

For the rest of the paper, we assume $d = m = 1$ and consider the example of (5.1) with $f(t, x, u) = -u$ and $g(t, x, u) = 0$:

$$-\partial_t u + u \partial_x u = 0 \text{ in } (0, T) \times \mathbb{R}, \quad u(T, \cdot) = u_T. \quad (5.15)$$

Observe that (5.15) can be formally written in conservative form, where it becomes the Burgers equation with flux $\frac{u^2}{2}$. When $d > 1$, it is not necessarily the case that (5.1) can be

written as a conservation law, and the product of $f(t, x, u)$ with ∇u cannot be understood by integrating by parts.

If the decreasing function u_T is Lipschitz continuous, then the system of characteristics (5.2) can be solved in some maximal time interval $[T - \tau, T]$ depending on the Lipschitz constant for u_T . This gives rise to a Lipschitz continuous solution of (5.15), which can easily be checked to be a viscosity solution in the sense of Definition 5.3 and is therefore unique. However, for $t < \tau$, the system of characteristics fails to be solvable on $[t, T]$, due to the formation of shocks. This is in contrast to the case where u_T is Lipschitz and *increasing*, in which case (5.15) has a Lipschitz continuous solution on $(-\infty, T]$. We therefore see that the situation where a continuous solution exists is not typical on an arbitrary time horizon, even if the function u_T is smooth.

We therefore study in this subsection a simple example of a decreasing and discontinuous terminal data, namely

$$u_T(x) = \mathbf{1}\{x \leq 0\}. \quad (5.16)$$

Viewed as a scalar conservation law with flux $\frac{u^2}{2}$, (5.15) is a Riemann problem whose solvability is resolved with the theory of entropy solutions. Indeed, the unique entropy solution is given by

$$u(t, x) = \mathbf{1}\left\{x \leq \frac{T-t}{2}\right\}, \quad (5.17)$$

describing a shock wave moving with constant speed $\frac{1}{2}$ (in reverse time), the constant $\frac{1}{2}$ being uniquely determined from the Rankine–Hugoniot condition.

5.4.1. Nonuniqueness of discontinuous solutions. When viewed as a nonlinear transport equation, there is a strong failure of uniqueness for this problem.

Proposition 5.4. *Let $c: [0, T] \rightarrow \mathbb{R}$ satisfy $c(T) = 0$ and $-c' \in [0, 1]$ a.e. Then*

$$u_c(t, x) := \mathbf{1}\{x \leq c(t)\}$$

is a solution, in the sense of Definition 5.1, of (5.15) with terminal data (5.16).

Proof. Define $u_c^\varepsilon = u_c * \rho^\varepsilon$. Then

$$u_c^\varepsilon(t, x) = \int_{-\infty}^{c(t)} \rho^\varepsilon(x - y) dy,$$

and so

$$\partial_t u_c^\varepsilon(t, x) = \rho^\varepsilon(x - c(t))c'(t) \quad \text{and} \quad \partial_x u_c^\varepsilon(t, x) = -\rho^\varepsilon(x - c(t)).$$

Thus,

$$-\partial_t u_c^\varepsilon(t, x) + u_c(t, x)\partial_x u_c^\varepsilon(t, x) = \rho^\varepsilon(x - c(t))[-c'(t) - \mathbf{1}\{x \leq c(t)\}].$$

Recall that $\text{supp } \rho^\varepsilon \in [-2\varepsilon, 0]$, and so the above expression is nonzero only if $x \leq c(t)$. In that case, because $-c' \leq 1$,

$$-c'(t) - \mathbf{1}\{x \leq c(t)\} = -c'(t) - 1 \leq 0,$$

and we conclude that u_c^ε is a subsolution. The proof of the supersolution property is similar, and uses the fact that $-c' \geq 0$. ■

Remark 5.4. Note that u_c is continuous a.e. in $[0, T] \times \mathbb{R}^d$, and, therefore, by Lemma 5.3, is also a viscosity solution in the sense of Definition 5.3.

The system of characteristics (5.2) becomes, for $(t, x) \in [0, T] \times \mathbb{R}^d$,

$$\begin{cases} -\partial_s U_{s,t}(x) = 0, & U_{T,t}(x) = \mathbf{1}\{X_{T,t}(x) \leq 0\}, \\ \partial_s X_{s,t}(x) = -U_{s,t}(x), & X_{t,t}(x) = x. \end{cases} \quad (5.18)$$

Recall that the system, and, in particular, the equation for X , is viewed as a differential inclusion, where, for all $s \in [t, T]$,

$$U_{s,t}(x) = U_{T,t}(x) = \begin{cases} \{1\} & \text{if } X_{T,t}(x) < 0, \\ \{0\} & \text{if } X_{T,t}(x) > 0, \text{ and} \\ [0, 1] & \text{if } X_{T,t}(x) = 0. \end{cases}$$

By Proposition 5.2, each solution u_c corresponds to a solution (X^c, U^c) of (5.18), which we can compute explicitly: namely, for $x \in \mathbb{R}$ and $0 \leq s \leq t \leq T$,

$$U_{s,t}^c(x) = u_c(t, x) = \mathbf{1}\{x \leq c(t)\}$$

and

$$X_{s,t}^c(x) = \begin{cases} x - s + t, & x \leq c(t), \\ x - c(t) \frac{s-t}{T-t}, & x = c(t), \\ x, & x \geq c(t). \end{cases} \quad (5.19)$$

We note that, for any solution (U, X) , we must have $X_{s,t}(x) = x - s + t$ for $x < 0$ and $X_{s,t}(x) = x$ for $x \geq T - t$. However, for $x \in [0, T - t]$, there is ambiguity in the speed at which the $X_{s,t}(x)$ travels: it can move with speed -1 , 0 , or anything in between, where in the latter case the characteristic is constrained to end at $X(T, t, x) = 0$. The precise value $x \in [0, 1]$ for which $X_{T,t}(x) = 0$ thus encodes the choice of the shock-wave speed $c(t)$ in the definition of u_c .

5.4.2. Stochastic selection. An important feature in the theory of entropy solutions of scalar conservation laws is the stability under regularizations, and in particular under vanishing viscosity limits. In the above context, the entropy solution (5.17) of (5.15) arises as the strong limit in $C([0, T], L_{\text{loc}}^1(\mathbb{R}^d))$, as $\varepsilon \rightarrow 0$, of the unique smooth solution u^ε of

$$-\partial_t u^\varepsilon + u^\varepsilon \partial_x u^\varepsilon = \varepsilon \partial_x^2 u^\varepsilon, \quad u^\varepsilon(T, \cdot) = u_T^\varepsilon,$$

where

$$u_T^\varepsilon: \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ is smooth and } \lim_{\varepsilon \rightarrow 0} u_T^\varepsilon = \mathbf{1}_{(-\infty, 0)} \text{ in } L_{\text{loc}}^1. \quad (5.20)$$

By contrast, we show here that any solution u_c can arise as a limit from suitably regularized equations.

Theorem 5.4. *Fix any $c \in W^{2,1}([0, T], \mathbb{R})$ satisfying $c(T) = 0$ and $-c' \in (0, 1)$. Then there exists $\theta_c \in L^1([0, T])$ such that, if u_T^ε is as in (5.20) and u^ε is the unique classical solution of*

$$-\partial_t u^\varepsilon + u^\varepsilon \partial_x u^\varepsilon = \varepsilon (\partial_x^2 u^\varepsilon + \theta_c(t) |\partial_x u^\varepsilon|^2), \quad u^\varepsilon(T, \cdot) = u_T^\varepsilon, \quad (5.21)$$

then, for all $1 \leq p < \infty$, as $\varepsilon \rightarrow 0$, $u^\varepsilon \rightarrow u_c$ strongly in $C([0, T], L_{\text{loc}}^p)$.

Proof. For $\theta: [0, T] \rightarrow \mathbb{R}$ to be determined, define, for $t \in [0, T]$,

$$f(t, v) = \begin{cases} \frac{\log(\theta(t)v + 1)}{\theta(t)} & \text{if } \theta(t) \neq 0 \text{ and } \theta(t)v > -1, \\ v & \text{if } \theta(t) = 0, \end{cases}$$

and note that

$$f(t, \cdot)^{-1}(u) = \frac{e^{\theta(t)u} - 1}{\theta(t)}.$$

For $\varepsilon > 0$, let v^ε be the classical solution of

$$\begin{cases} -\partial_t v^\varepsilon + f(t, v^\varepsilon) v_x^\varepsilon = \frac{\partial_t f(t, v^\varepsilon)}{\partial_v f(t, v^\varepsilon)} + \varepsilon \partial_x^2 v^\varepsilon & \text{in } [0, T] \times \mathbb{R} \text{ and} \\ v^\varepsilon(T, \cdot) = f(T, \cdot)^{-1}(u_T^\varepsilon). \end{cases}$$

By the maximum principle, it is easily checked that the values of v^ε fall within the domain of f and its derivatives, and the solution of (5.21) is given exactly by $u^\varepsilon(t, x) = f(t, v^\varepsilon(t, x))$.

Standard stability results [40] yield that, as $\varepsilon \rightarrow 0$,

$$v^\varepsilon \rightarrow v \quad \text{strongly in } C([0, T], L_{\text{loc}}^p)$$

for all $p \in [1, \infty)$, where v is the unique entropy solution of

$$\begin{cases} -\partial_t v + f(t, v) \partial_x v = \frac{\partial_t f(t, v)}{\partial_v f(t, v)} & \text{in } [0, T] \times \mathbb{R} \text{ and} \\ v(T, \cdot) = f^{-1}(T, \cdot)(1) \mathbf{1}_{x < 0} + f^{-1}(T, \cdot)(0) \mathbf{1}_{x > 0}. \end{cases} \quad (5.22)$$

We then set

$$F(t, v) = \frac{(\theta(t)v + 1) \log(\theta(t)v + 1) - \theta(t)v}{\theta(t)^2},$$

which satisfies $\partial_v F(t, v) = f(t, v)$, so that equation (5.22) is equivalently written as

$$\begin{cases} -\partial_t v + \partial_x F(t, v) = \frac{\partial_t f(t, v)}{\partial_v f(t, v)} & \text{in } [0, T] \times \mathbb{R} \text{ and} \\ v(T, \cdot) = f^{-1}(T, \cdot)(1)\mathbf{1}_{x < 0} + f^{-1}(T, \cdot)(0)\mathbf{1}_{x > 0}. \end{cases} \quad (5.23)$$

The unique entropy solution v of the conservative equation (5.23) is then given by

$$v(t, x) = \begin{cases} v^-(t), & x < c(t), \\ v^+(t), & x > c(t), \end{cases}$$

where, for $0 \leq t \leq T$, v_t^+ and v_t^- are defined by

$$-\partial_t v_t^\pm = \frac{\partial_t f(t, v_t^\pm)}{\partial_v f(t, v_t^\pm)}, \quad t \in [0, T], \quad v_T^- = f^{-1}(T, \cdot)(0), \quad v_T^+ = f^{-1}(T, \cdot)(1),$$

and $c(t)$ is determined by the Rankine–Hugoniot condition

$$c(T) = 0 \quad \text{and} \quad -c'(t) = \frac{F(t, v^-(t)) - F(t, v^+(t))}{v^+(t) - v^-(t)}.$$

Observe that $t \mapsto f(t, v^\pm(t))$ is constant, and therefore

$$-c'(t) = \frac{F(t, f(t, \cdot)^{-1}(1)) - F(t, f(t, \cdot)^{-1}(0))}{f(t, \cdot)^{-1}(1) - f(t, \cdot)^{-1}(0)} = C(\theta(t)),$$

where, for $\theta \in \mathbb{R}$,

$$C(\theta) = \frac{\theta e^\theta - e^\theta - 1}{\theta(e^\theta - 1)}.$$

We observe that $C' > 0$, $\lim_{\theta \rightarrow -\infty} C(\theta) = 0$, and $\lim_{\theta \rightarrow +\infty} C(\theta) = 1$. Thus, letting $c: [0, T] \rightarrow \mathbb{R}$ with $-c' \in (0, 1)$ be as in the statement of the theorem, we finally choose $\theta_c(t) := C^{-1}(-c'(t))$, so that u^ε converges strongly in $C([0, T], L_{\text{loc}}^p)$, $p \in [1, \infty)$, to $f^{-1}(t, \cdot)(v(t, x)) = u_c(t, x)$. ■

Remark 5.5. The restrictions that $c'' \in L^1$ and c' lie strictly within $(-1, 0)$ are put in place to ensure that $\theta_c \in W^{1,1}$ and $-\infty < \theta_c < \infty$, so that the equation for v^ε is well posed. Achieving speeds c that are only Lipschitz, and where c' is allowed to be either 0 or 1, is possible by letting $\theta_c = \theta_c^\varepsilon$ depend suitably on ε .

The selection principle in Theorem 5.4 can be recast in terms of the nonunique generalized characteristics problem (5.18). Namely, we fix a filtered probability space $(\Omega, \mathbb{F} = (\mathbb{F}_t)_{t \in [0, T]}, \mathbb{P})$ with a complete, right-continuous filtration \mathbb{F} and a Wiener process $W: [0, T] \times \Omega \rightarrow \mathbb{R}$ progressively measurable with respect to \mathbb{F} , and, for $\varepsilon > 0$ and $(t, x) \in [0, T] \times \mathbb{R}$, we consider the forward-backward SDE (FBSDE) on the interval $s \in [t, T]$:

$$\begin{cases} -d_s U_{s,t}^\varepsilon(x) = Z_{s,t}^\varepsilon(x) dW_s - \frac{1}{2} \theta_c(s) Z_{s,t}^\varepsilon(x)^2 ds, & U_{T,t}^\varepsilon(x) = u_T^\varepsilon(X_{T,t}^\varepsilon(x)), \\ d_s X_{s,t}^\varepsilon(x) = -U_{s,t}^\varepsilon(x) ds + \sqrt{2\varepsilon} dW_s, & X_{t,t}^\varepsilon(x) = x. \end{cases} \quad (5.24)$$

Theorem 5.5. *Let c and θ_c be as in Theorem 5.4. For every $\varepsilon > 0$ and $(t, x) \in [0, T] \times \mathbb{R}$, there exists a unique strong solution $(X_{s,t}^\varepsilon(x), U_{s,t}^\varepsilon(x), Z_{s,t}^\varepsilon(x))$ to the FBSDE (5.24). Moreover, with probability 1, as $\varepsilon \rightarrow 0$, $X_{s,t}^\varepsilon \rightarrow X_{s,t}^c$ in $C([t, T], L_{\text{loc}}^p(\mathbb{R}))$ for all $p \in [1, \infty)$, where X^c is as in (5.19).*

Proof. The existence of a unique solution is a consequence of the nondegeneracy of the noise; see [52]. Moreover, if u^ε solves (5.21), then

$$U_{s,t}^\varepsilon(x) = u^\varepsilon(s, X_{s,t}^\varepsilon(x)) \quad \text{and} \quad Z_{s,t}^\varepsilon(x) = \sqrt{2\varepsilon}(\partial_x u^\varepsilon)(s, X_{s,t}^\varepsilon(x)).$$

It follows that $X_{s,t}^\varepsilon$ is the solution of the SDE (3.7) with $b(t, x) = -u^\varepsilon(t, x)$, and so, by Theorems 3.3 and 5.4, as $\varepsilon \rightarrow 0$ with probability 1, $X_{s,t}^\varepsilon$ converges in $C([t, T], L_{\text{loc}}^p(\mathbb{R}))$ to the regular Lagrangian flow corresponding to the drift $-u_c$. This is exactly X^c , and we conclude. ■

Funding. B. Seeger acknowledges support from the National Science Foundation under award numbers DMS-1840314 and DMS-2307610.

References

- [1] S. Ahuja, [Wellposedness of mean field games with common noise under a weak monotonicity condition](#). *SIAM J. Control Optim.* **54** (2016), no. 1, 30–48 Zbl 1327.93403 MR 3439756
- [2] L. Ambrosio, [Transport equation and Cauchy problem for \$BV\$ vector fields](#). *Invent. Math.* **158** (2004), no. 2, 227–260 Zbl 1075.35087 MR 2096794
- [3] L. Ambrosio, [Transport equation and Cauchy problem for non-smooth vector fields](#). In *Calculus of variations and nonlinear partial differential equations*, pp. 1–41, Lecture Notes in Math. 1927, Springer, Berlin, 2008 MR 2408257
- [4] L. Ambrosio, [Well posedness of ODE’s and continuity equations with nonsmooth vector fields, and applications](#). *Rev. Mat. Complut.* **30** (2017), no. 3, 427–450 Zbl 1380.37039 MR 3690748
- [5] L. Ambrosio and G. Crippa, [Continuity equations and ODE flows with non-smooth velocity](#). *Proc. Roy. Soc. Edinburgh Sect. A* **144** (2014), no. 6, 1191–1244 Zbl 1358.37046 MR 3283066
- [6] L. Ambrosio and D. Trevisan, [Lecture notes on the DiPerna–Lions theory in abstract measure spaces](#). *Ann. Fac. Sci. Toulouse Math. (6)* **26** (2017), no. 4, 729–766 Zbl 1402.35003 MR 3746641
- [7] C. Arzelà, [Sulle funzioni di due variabili a variazione limitata](#). *Bologna Rend.* 1904/05. 7 S. (1904,1905). Zbl 36.0491.02 Icuni teoremi sulle funzioni semicontinue, e sulle funzioni di una variabile, limiti di funzioni di due variabili reali
- [8] C. Bertucci, [Monotone solutions for mean field games master equations: Finite state space and optimal stopping](#). *J. Éc. polytech. Math.* **8** (2021), 1099–1132 Zbl 1472.35394 MR 4275225
- [9] C. Bertucci, [Monotone solutions for mean field games master equations: Continuous state space and common noise](#). *Comm. Partial Differential Equations* **48** (2023), no. 10–12, 1245–1285 Zbl 1531.91027 MR 4687277

- [10] C. Bertucci, J.-M. Lasry, and P.-L. Lions, [Some remarks on mean field games](#). *Comm. Partial Differential Equations* **44** (2019), no. 3, 205–227 Zbl [1411.91100](#) MR [3941633](#)
- [11] C. Bertucci, J.-M. Lasry, and P.-L. Lions, [On Lipschitz solutions of mean field games master equations](#). *J. Funct. Anal.* **287** (2024), no. 5, article no. 110486 MR [4749404](#)
- [12] F. Bouchut and F. James, [One-dimensional transport equations with discontinuous coefficients](#). *Nonlinear Anal.* **32** (1998), no. 7, 891–933 Zbl [0989.35130](#) MR [1618393](#)
- [13] F. Bouchut, F. James, and S. Mancini, [Uniqueness and weak stability for multi-dimensional transport equations with one-sided Lipschitz coefficient](#). *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **4** (2005), no. 1, 1–25 Zbl [1170.35363](#) MR [2165401](#)
- [14] S. Breneis, *Functions of bounded variation in one and multiple dimensions*. Master’s thesis, Johannes Kepler Universität Linz, 2020
- [15] J. Busca and B. Sirakov, [Harnack type estimates for nonlinear elliptic systems and applications](#). *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **21** (2004), no. 5, 543–590 Zbl [1127.35332](#) MR [2086750](#)
- [16] I. Capuzzo Dolcetta and A. Vitolo, Weak maximum principle for cooperative systems: The degenerate elliptic case. *J. Convex Anal.* **28** (2021), no. 2, 495–508 Zbl [1471.35120](#) MR [4245878](#)
- [17] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, [The master equation and the convergence problem in mean field games](#). *Ann. of Math. Stud.* 201, Princeton University Press, Princeton, NJ, 2019 Zbl [1430.91002](#) MR [3967062](#)
- [18] P. Cardaliaguet and P. Souganidis, [Monotone solutions of the master equation for mean field games with idiosyncratic noise](#). *SIAM J. Math. Anal.* **54** (2022), no. 4, 4198–4237 Zbl [1497.35471](#) MR [4451309](#)
- [19] P. Cardaliaguet and P. E. Souganidis, [On first order mean field game systems with a common noise](#). *Ann. Appl. Probab.* **32** (2022), no. 3, 2289–2326 Zbl [1498.91040](#) MR [4430014](#)
- [20] R. Carmona and F. Delarue, [Probabilistic theory of mean field games with applications. II](#). *Probab. Theory Stoch. Model.* 84, Springer, Cham, 2018 Zbl [1422.91015](#) MR [3753660](#)
- [21] A. Cecchin and F. Delarue, [Selection by vanishing common noise for potential finite state mean field games](#). *Comm. Partial Differential Equations* **47** (2022), no. 1, 89–168 Zbl [1490.91018](#) MR [4373880](#)
- [22] N. Champagnat and P.-E. Jabin, [Strong solutions to stochastic differential equations with rough coefficients](#). *Ann. Probab.* **46** (2018), no. 3, 1498–1541 Zbl [1451.60091](#) MR [3785594](#)
- [23] E. D. Conway, [Generalized solutions of linear differential equations with discontinuous coefficients and the uniqueness question for multidimensional quasilinear conservation laws](#). *J. Math. Anal. Appl.* **18** (1967), 238–251 Zbl [0163.12103](#) MR [0206474](#)
- [24] M. G. Crandall, H. Ishii, and P.-L. Lions, [User’s guide to viscosity solutions of second order partial differential equations](#). *Bull. Amer. Math. Soc. (N.S.)* **27** (1992), no. 1, 1–67 Zbl [0755.35015](#) MR [1118699](#)
- [25] A. M. Davie, [Uniqueness of solutions of stochastic differential equations](#). *Int. Math. Res. Not. IMRN* (2007), no. 24, article no. rnm124 Zbl [1139.60028](#) MR [2377011](#)
- [26] F. Delarue and R. Foguen Tchuendom, [Selection of equilibria in a linear quadratic mean-field game](#). *Stochastic Process. Appl.* **130** (2020), no. 2, 1000–1040 Zbl [1471.91029](#) MR [4046528](#)
- [27] J. Dianetti, Strong solutions to submodular mean field games with common noise and related McKean–Vlasov FBSDEs. Center for Mathematical Economics Working Papers 674, Center for Mathematical Economics, Bielefeld University

- [28] J. Dianetti, G. Ferrari, M. Fischer, and M. Nendel, [Submodular mean field games: Existence and approximation of solutions](#). *Ann. Appl. Probab.* **31** (2021), no. 6, 2538–2566
Zbl [1485.91020](#) MR [4350967](#)
- [29] J. Dianetti, G. Ferrari, M. Fischer, and M. Nendel, [A unifying framework for submodular mean field games](#). *Math. Oper. Res.* **48** (2023), no. 3, 1679–1710 Zbl [1533.91042](#) MR [4702114](#)
- [30] R. J. DiPerna and P.-L. Lions, [Ordinary differential equations, transport theory and Sobolev spaces](#). *Invent. Math.* **98** (1989), no. 3, 511–547 Zbl [0696.34049](#) MR [1022305](#)
- [31] A. Figalli, [Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients](#). *J. Funct. Anal.* **254** (2008), no. 1, 109–153 Zbl [1169.60010](#) MR [2375067](#)
- [32] A. F. Filippov, [Differential equations with discontinuous right-hand side](#). *Mat. Sb. (N.S.)* **51(93)** (1960), 99–128 Zbl [0138.32204](#) MR [0114016](#)
- [33] W. Gangbo and A. R. Mészáros, [Global well-posedness of master equations for deterministic displacement convex potential mean field games](#). *Comm. Pure Appl. Math.* **75** (2022), no. 12, 2685–2801 Zbl [1510.35342](#) MR [4509653](#)
- [34] W. Gangbo, A. R. Mészáros, C. Mou, and J. Zhang, [Mean field games master equations with nonseparable Hamiltonians and displacement monotonicity](#). *Ann. Probab.* **50** (2022), no. 6, 2178–2217 Zbl [1501.35403](#) MR [4499277](#)
- [35] D. A. Gomes, J. Mohr, and R. R. Souza, [Continuous time finite state mean field games](#). *Appl. Math. Optim.* **68** (2013), no. 1, 99–143 Zbl [1283.91016](#) MR [3072242](#)
- [36] P. J. Graber and A. R. Mészáros, [On monotonicity conditions for mean field games](#). *J. Funct. Anal.* **285** (2023), no. 9, article no. 110095 Zbl [1519.91034](#) MR [4620724](#)
- [37] P. J. Graber and A. R. Mészáros, [On some mean field games and master equations through the lens of conservation laws](#). *Math. Ann.* (2024)
DOI <https://doi.org/10.1007/s00208-024-02859-z>
- [38] M. Grassin and D. Serre, [Existence de solutions globales et régulières aux équations d’Euler pour un gaz parfait isentropique](#). *C. R. Acad. Sci. Paris Sér. I Math.* **325** (1997), no. 7, 721–726
Zbl [0887.35125](#) MR [1483706](#)
- [39] H. Ishii, [Perron’s method for Hamilton–Jacobi equations](#). *Duke Math. J.* **55** (1987), no. 2, 369–384 Zbl [0697.35030](#) MR [0894587](#)
- [40] S. N. Kružkov, [First order quasilinear equations with several independent variables](#). *Mat. Sb. (N.S.)* **81(123)** (1970), 228–255 Zbl [0202.11203](#) MR [0267257](#)
- [41] J.-M. Lasry and P.-L. Lions, [Jeux à champ moyen. I. Le cas stationnaire](#). *C. R. Math. Acad. Sci. Paris* **343** (2006), no. 9, 619–625 Zbl [1153.91009](#) MR [2269875](#)
- [42] J.-M. Lasry and P.-L. Lions, [Jeux à champ moyen. II. Horizon fini et contrôle optimal](#). *C. R. Math. Acad. Sci. Paris* **343** (2006), no. 10, 679–684 Zbl [1153.91010](#) MR [2271747](#)
- [43] J.-M. Lasry and P.-L. Lions, [Mean field games](#). *Jpn. J. Math.* **2** (2007), no. 1, 229–260
Zbl [1156.91321](#) MR [2295621](#)
- [44] J.-M. Lasry, P.-L. Lions, and B. Seeger, [Dimension reduction techniques in deterministic mean field games](#). *Comm. Partial Differential Equations* **47** (2022), no. 4, 701–723
Zbl [1490.91019](#) MR [4417027](#)
- [45] C. Le Bris and P.-L. Lions, [Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients](#). *Comm. Partial Differential Equations* **33** (2008), no. 7–9, 1272–1317 Zbl [1157.35301](#) MR [2450159](#)
- [46] C. Le Bris and P.-L. Lions, [Parabolic equations with irregular data and related issues—Applications to stochastic differential equations](#). De Gruyter Ser. Appl. Numer. Math. 4, De Gruyter, Berlin, 2019 Zbl [07046231](#) MR [4319031](#)

- [47] P.-L. Lions, Jeux à champ moyen. 2011–12, course at Collège de France, <https://www.youtube.com/watch?v=vmqlduGczw4> visited on 3 July 2024
- [48] P.-L. Lions and B. Seeger, Transport equations and flows with one-sided Lipschitz velocity fields. 2023, arXiv:2306.13288v1
- [49] H. Liu and E. Tadmor, Spectral dynamics of the velocity gradient field in restricted flows. *Comm. Math. Phys.* **228** (2002), no. 3, 435–466 Zbl 1031.76006 MR 1918784
- [50] C. Mou and J. Zhang, Mean field game master equations with anti-monotonicity conditions. *J. Eur. Math. Soc. (JEMS)* (2024) DOI <https://doi.org/10.4171/jems/1455>
- [51] C. Mou and J. Zhang, Minimal solutions of master equations for extended mean field games. *J. Math. Pures Appl. (9)* **184** (2024), 190–217 Zbl 07827006 MR 4717265
- [52] E. Pardoux and S. Tang, Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory Related Fields* **114** (1999), no. 2, 123–150 Zbl 0943.60057 MR 1701517
- [53] G. Petrova and B. Popov, Linear transport equations with discontinuous coefficients. *Comm. Partial Differential Equations* **24** (1999), no. 9-10, 1849–1873 Zbl 0992.35104 MR 1708110
- [54] G. Petrova and B. Popov, Linear transport equations with μ -monotone coefficients. *J. Math. Anal. Appl.* **260** (2001), no. 2, 307–324 Zbl 0980.35003 MR 1845556
- [55] F. Poupaud and M. Rascle, Measure solutions to the linear multi-dimensional transport equation with non-smooth coefficients. *Comm. Partial Differential Equations* **22** (1997), no. 1-2, 337–358 Zbl 0882.35026 MR 1434148
- [56] W. Stepanoff, Über totale Differenzierbarkeit. *Math. Ann.* **90** (1923), no. 3-4, 318–320 Zbl 49.0183.01 MR 1512177
- [57] A. Tarski, A lattice-theoretical fixpoint theorem and its applications. *Pacific J. Math.* **5** (1955), 285–309 Zbl 0064.26004 MR 0074376
- [58] D. Trevisan, Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.* **21** (2016), article no. 22 Zbl 1336.60159 MR 3485364
- [59] A. J. Veretennikov, Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S.)* **111(153)** (1980), no. 3, 434–452 Zbl 0431.60061 MR 0568986

Received 28 July 2023; revised 11 March 2024; accepted 2 May 2024.

Pierre-Louis Lions

CEREMADE, Université Paris Dauphine, Pl. du Maréchal de Lattre de Tassigny, 75016 Paris;
Collège de France, 3 Rue d’Ulm, 75005 Paris, France; lions@ceremade.dauphine.fr

Benjamin Seeger

Department of Mathematics, University of Texas at Austin, 2515 Speedway, Austin, TX 78712,
USA; seeger@math.utexas.edu