Proceedings of the Royal Society of Edinburgh, **148A**, 1279–1296, *2018*
DOI:10.1017/S0308210518000173

Optimal stability estimates for continuity equations

Christian Seis*
Institut für Angewandte Mathematik, Universität Bonn,
Endenicher Allee 60, 53115 Bonn, Germany

(MS received 22 August 2016; accepted 1 December 2016)

This review paper is concerned with the stability analysis of the continuity equation in the DiPerna–Lions setting in which the advecting velocity field is Sobolev regular. Quantitative estimates for the equation were derived only recently, but optimality was not discussed. We revisit the results from our 2017 paper, compare the new estimates with previously known estimates for Lagrangian flows and demonstrate how these can be applied to produce optimal bounds in applications from physics, engineering and numerical analysis.

Keywords: continuity equation; Sobolev vector fields; stability estimates; Kantorovich–Rubinstein distance; Lagrangian flows

2010 *Mathematics subject classification:* Primary 35F10
Secondary 35B30; 37C10

1. Introduction

The linear continuity equation is one of the most elementary partial differential equations. It describes the conservative transport of a quantity by a vector field. We shall study this equation in a convex and bounded Lipschitz domain Ω in $\mathbb{R}^d$ and denote by $\rho(t,x) \in \mathbb{R}$ and $u(t,x) \in \mathbb{R}^d$ the quantity and the vector field, respectively. (All the results in this work can be extended to arbitrary Lipschitz domains, the periodic torus or all of $\mathbb{R}^d$ with suitable modifications.) For a given initial configuration $\bar\rho(x) \in \mathbb{R}$, the Cauchy problem for the continuity equation reads

$$\left.\begin{aligned} \partial_t \rho + \nabla\cdot(u\rho) = 0 &\quad \text{in } (0,\infty)\times\Omega,\\ \rho(0,\cdot) = \bar\rho &\quad \text{in } \Omega. \end{aligned}\right\} \tag{1.1}$$

If the vector field is tangential at the boundary of Ω, which we assume from here on, the quantity ρ is (formally) conserved by the flow:

$$\forall t > 0, \quad \int_\Omega \rho(t,x)\,\mathrm{d}x = \int_\Omega \bar\rho(x)\,\mathrm{d}x.$$

Despite its simplicity, the continuity equation plays an important role in fluid dynamics and the theory of conservation laws. In typical applications, ρ represents

*Present address: Institut für Analysis und Numerik, Universität Münster, Orléans-Ring 10, 48149 Münster, Germany (seis@wwu.de).

mass or number density, temperature, energy or phase indicator. In the following, we shall frequently refer to ρ as a (possibly negative) mass density, or simply a density. Note that the vector field has dimensions of length divided by time, and we shall accordingly often refer to u as a velocity field. In fluids applications, u is the velocity of the fluid.

There is a close link between the partial differential equation (PDE) (1.1) and the following ordinary differential equation (ODE):

$$\left.\begin{aligned}\partial_t\phi(t,x) &= u(t,\phi(t,x)),\\ \phi(0,x) &= x.\end{aligned}\right\} \tag{1.2}$$

While the PDE represents the Eulerian specification of the flow field, i.e. the description of the dynamics at a fixed location and time, the ODE is the Lagrangian specification, which traces single particles through space and time. The two specifications are in fact equivalent: in the smooth setting, the solution to the continuity equation takes the nice form

$$\rho(t,\phi(t,x))\det\nabla\phi(t,x) = \bar\rho(x), \quad \text{or simply } \rho(t,\cdot) = (\phi(t,\cdot))_{\#}\bar\rho, \tag{1.3}$$

i.e. ρ is the push-forward of $\bar\rho$ by the flow ϕ, and a similar formula holds in the non-smooth setting, as long as (1.1) and (1.2) are well-posed. This superposition principle is reviewed in [3, § 3]. Note that, for any fixed time t, the mapping $\phi(t,\cdot)$ is a diffeomorphism on Ω, whose existence is obtained by the classical Picard–Lindelöf theorem, and $\det\nabla\phi(t,\cdot)$ is the Jacobian determinant, denoted by $J\phi(t,\cdot)$ in the following. The solution ϕ of (1.2) is called the *flow* of the vector field u.

Outside the smooth setting, well-posedness theory for both the PDE (1.1) and the ODE (1.2) is more challenging. We focus on the continuity equation from here on, and we start with a suitable concept of generalized solutions in the case where $\bar\rho \in L^q(\Omega)$ for some $q \in [1,\infty]$. We call ρ a *distributional solution* to the continuity equation (1.1) in the time interval $[0,T]$ if it conserves the integrability class of the initial datum $\rho \in L^\infty((0,T);L^q(\Omega))$ and satisfies

$$-\int_0^T\int_\Omega(\partial_t\zeta + u\cdot\nabla\zeta)\rho\,\mathrm{d}x\,\mathrm{d}t = \int_\Omega\bar\rho\zeta(0,\cdot)\,\mathrm{d}x$$

for any $\zeta \in C_c^\infty([0,T)\times\Omega)$. This distributional formulation is reasonable if $u \in L^1((0,T);L^p(\Omega))$ with $1/p + 1/q = 1$.

In order to prove the existence of distributional solutions, we shall impose a condition on the compressibility of the vector field: if u is weakly compressible, i.e.

$$(\nabla\cdot u)^- \in L^1((0,T);L^\infty(\Omega)), \tag{1.4}$$

then the existence is easily obtained by approximation with smooth functions. Here we have used the superscript minus sign to denote the negative part of the divergence.

The questions of uniqueness and continuous dependence on the initial data are more delicate, and were first answered positively by DiPerna and Lions in their groundbreaking paper [20]. Their theory is based on a new solution concept: the theory of renormalized solutions. A *renormalized solution* is a distributional solution

ρ with the property that, for any bounded function $\beta \in C^1(\mathbb{R})$ with bounded derivatives and $\beta(0) = 0$, the composition $\beta(\rho)$ satisfies the continuity equation with source

$$\partial_t \beta(\rho) + \nabla \cdot (u\beta(\rho)) = (\nabla \cdot u)(\beta(\rho) - \rho\beta'(\rho))$$

in the sense of distributions. In fact, under the additional assumption that u is in $L^1((0,T); W^{1,p}(\Omega))$, DiPerna and Lions showed that distributional solutions *are* renormalized solutions. (Their result has been extended by Ambrosio [2] to vector fields with bounded variation (BV) regularity, and, more recently, by Crippa *et al*. [12] to the case where the velocity gradient is given by a singular integral of an L^1 function.) The advantage of this solution concept is apparent: by choosing $\beta(s)$ as a suitable approximation of $|s|^q$, we obtain, by integration over Ω, that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega |\rho|^q \,\mathrm{d}x = -(q-1) \int_\Omega (\nabla \cdot u)|\rho|^q \,\mathrm{d}x \leqslant (q-1)\|(\nabla \cdot u)^-\|_{L^\infty(\Omega)} \int_\Omega |\rho|^q \,\mathrm{d}x,$$

and thus, with the help of the Gronwall lemma,

$$\sup_{(0,T)} \|\rho\|_{L^q(\Omega)} \leqslant \exp\left(\int_0^T \|(\nabla \cdot u)^-\|_{L^\infty(\Omega)} \,\mathrm{d}t \right)^{1-1/q} \|\bar{\rho}\|_{L^q(\Omega)}. \tag{1.5}$$

By the linearity of the continuity equation, this estimate implies both uniqueness and continuous dependence on the initial data.

In addition to proving well-posedness, DiPerna and Lions studied stability under approximations of the vector fields and under diffusive perturbations of the equation. (Note that this gives two different ways of regularizing the PDE.) While qualitative stability is obtained easily via renormalization, the theory fails to provide quantitative stability estimates that capture the rate of convergence of approximate or perturbative solutions to the original equation. Such estimates were recently developed in [38].

The aim of the present paper is to revisit the stability estimates from [38] and to reformulate them in a new and optimal way. We shall mainly focus on two aspects. Firstly, our intention is to compare the new results with earlier achievements in the theory of Lagrangian flows [11] (which, in fact, strongly inspired the estimates in [38]). By doing so, we hope to convince the reader that the quantities considered in [38] appear naturally in the context of continuity equations. Secondly, we shall present applications of the estimates that allow us to compute optimal convergence rates in examples of approximate vector fields, zero-diffusivity limits, fluid mixing and numerical upwind schemes. The last two examples are taken from [36, 37]. We include these results in order to demonstrate the strength of the estimates from [38], and to underline the intrinsic connection between the respective works (in particular [37,38]). The first example partly extends the recent analysis in [15].

We conclude this introduction by noting that, as a by-product of the stability estimates, in [38], a new proof of uniqueness is given for (1.1). This new proof does not rely on the theory of renormalized solutions but is based solely on the distributional formulation of the equation. In a way, the theory in [38] is the PDE counterpart of the quantitative theory for Lagrangian flows developed by Crippa and De Lellis in [11]. In fact, some of the key estimates were successively transferred from [11] to [38].

Notation In the following, we shall use the shorter notation L^r for the Lebesgue space $L^r(\Omega)$, and, similarly, $L^r(L^s)$ for $L^r((0,T);L^s(\Omega))$. Further function spaces such as $L^1(W^{1,p})$ are defined analogously.

We shall omit the domain of integration in the spatial integrals for notational convenience. For instance, we write $\int \cdot \, \mathrm{d}x$ for $\int_\Omega \cdot \, \mathrm{d}x$.

We use the sloppy notation $a \lesssim b$ if $a \leqslant Cb$ for some constant C that may depend only on the dimension d, the domain Ω or the Sobolev exponent p. We write $a \lesssim_{r_1,\dots,r_n} b$ if C also depends on the quantities $r_1,\dots,r_n$. Finally, we shall sometimes use the notation $a \sim b$ if $a \lesssim b$ and $b \lesssim a$.

2. Stability estimates for Lagrangian flows

To motivate our new perspective on the results from [38], we start by recalling some facts from the theory of particles moving in a weakly compressible fluid. The trajectory of a particle moving with the flow is given by the solution of the ODE (1.2). In the classical setting, when the advecting velocity field u is smooth, or at least Lipschitz continuous in the spatial variable, the existence and uniqueness of a solution are given by the Picard–Lindelöf theorem. The Lipschitz regularity also yields simple estimates on the distance between particle trajectories at any time during the evolution. Indeed, as a consequence of the elementary computation

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} |\phi(t,x) - \phi(t,y)| \right| \leqslant |u(t,\phi(t,x)) - u(t,\phi(t,y))| \leqslant \|\nabla u\|_{L^\infty} |\phi(t,x) - \phi(t,y)|$$

and the Gronwall lemma, we easily derive the estimate

$$\exp\left(-\int_0^t \|\nabla u\|_{L^\infty} \, \mathrm{d}t \right) \leqslant \frac{|\phi(t,x) - \phi(t,y)|}{|x-y|} \leqslant \exp\left(\int_0^t \|\nabla u\|_{L^\infty} \, \mathrm{d}t \right). \tag{2.1}$$

Here we have used Rademacher's identification of Lipschitz functions with the Sobolev class $W^{1,\infty}$, so that $\|\nabla u\|_{L^\infty}$ is the Lipschitz constant of u. This estimate illustrates the well-known fact that two particles transported by the flow can neither converge more slowly nor diverge more quickly than exponentially in time.

This classical result can equivalently be rewritten as

$$-\int_0^t \|\nabla u\|_{L^\infty} \, \mathrm{d}t \leqslant \log\left(\frac{|\phi(t,x) - \phi(t,y)|}{|x-y|} \right) \leqslant \int_0^t \|\nabla u\|_{L^\infty} \, \mathrm{d}t, \tag{2.2}$$

which shows that the velocity gradient controls the logarithmic relative distance between two particles. Here 'relative distance' refers to the actual distance between particles relative to their initial distance. We shall see in the following that it is the latter perspective rather than the classical one (2.1) that allows a generalization to the case of flows for less regular vector fields and to the Eulerian setting.

Note that both the estimates (2.1) and (2.2) contain some information on the regularity of the flow: the flow itself is spatially Lipschitz uniformly in time, with the Lipschitz constant depending on the gradient of u.

Instead of tracing the distance between two different particles in a fluid, we can similarly study the distance between trajectories corresponding to a particle transported by different vector fields: if ϕ and ϕ_k denote the flows associated via

(1.2) to the vector fields u and u_k, respectively, where u_k may be thought of as a Lipschitz continuous perturbation of u, then a computation similar to that above yields the estimate

$$\log\left(\frac{|\phi(t,x)-\phi_k(t,x)|}{\delta}+1\right) \leqslant \int_0^t \|\nabla u\|_{L^\infty}\,\mathrm{d}t + \frac{1}{\delta}\int_0^t \|u-u_k\|_{L^\infty}\,\mathrm{d}t \tag{2.3}$$

for any $\delta > 0$. Thus, by choosing $\delta = \delta_k(t) = \int_0^t \|u-u_k\|_{L^\infty}\,\mathrm{d}t$, we see that

$$\log\left(\frac{|\phi(t,x)-\phi_k(t,x)|}{\delta_k(t)}+1\right) \leqslant \int_0^t \|\nabla u\|_{L^\infty}\,\mathrm{d}t + 1, \tag{2.4}$$

so that, as before, the velocity gradient controls the logarithmic relative distance between particles moving with two different flows. Observe that $\delta_k(t)$ scales like a length, and can thus be interpreted as the (maximal) distance between the velocity fields. Hence, we control the distance between particles relative to the distance between vector fields.

Inequality (2.4) is an estimate on the rate of convergence of trajectories associated with the vector fields u and u_k, if the approximating vector field u_k converges to u in the sense that $\delta_k(t) \to 0$. Inequality (2.4) then shows that the particle trajectories approach each other with a rate of at least $\delta_k(t)$.

Note that (2.3) also implies uniqueness of (1.2) when the existence of a solution to the ODE is known. Indeed, if $u = u_k$ is spatially Lipschitz and ϕ and ϕ_k are two solutions to (1.2), the right-hand side of (2.3) is bounded independently of δ. Hence, choosing δ arbitrarily small, we see that ϕ and ϕ_k must be identical.

Outside of the smooth setting, the notion of flows for vector fields has to be appropriately generalized. A common generalization is the notion of regular Lagrangian flows that are well defined if u is merely Sobolev (or even BV) regular in the spatial variable and weakly compressible [2, 11, 20]. The latter is expressed by the requirement that

$$-\infty < \nabla\cdot u(t,x) \quad \text{for almost every } (t,x) \in (0,T)\times\Omega$$

(cf. (1.4)), which in turn implies that the Jacobi determinant is bounded from below:

$$J\phi(t,x) = \exp\left(\int_0^t \nabla\cdot u(t,\phi(t,x))\right)\mathrm{d}t) \geqslant \exp\left(-\int_0^T \|(\nabla\cdot u)^-\|_{L^\infty}\,\mathrm{d}t\right) =: \Lambda. \tag{2.5}$$

The weak compressibility condition excludes the possibility of infinitely strong sinks in which particles collide in finite time.

The existence, uniqueness and stability of regular Lagrangian flows were established in the case of vector fields with spatial Sobolev regularity (under the assumption that the divergence is uniformly bounded) by DiPerna and Lions in their seminal paper [20]. This theory has been substantially extended to BV vector fields by Ambrosio [2]. We refer the interested reader to [3, 13, 14] for more details and further references, and remark in addition that a comprehensive analysis of the Jacobian is given in [10].

Interestingly, DiPerna and Lions's theory for the ODE (1.2) is built on a well-posedness theory for the associated transport (cf. (3.9)) and continuity equations, that is, on the Eulerian (and thus PDE) perspective on particle dynamics. The drawback of the qualitative theory is that no quantitative estimates can be provided. Stability estimates of the type (2.4) in the DiPerna–Lions setting were derived later by Crippa and De Lellis [11]. These are of the form

$$\fint \log\left(\frac{|\phi(t,x)-\phi_k(t,x)|}{\delta_k(t)}+1\right)\mathrm{d}x \lesssim_{\Lambda} \int_0^t \|\nabla u\|_{L^p}\,\mathrm{d}t + 1, \tag{2.6}$$

where we now have

$$\delta_k(t) = \int_0^t \|u-u_k\|_{L^p}\,\mathrm{d}t.$$

We decided to work with averaged spatial integrals in all formulae in this paper. Therefore, $\int_0^t \|\nabla u\|_{L^p}\,\mathrm{d}t$ is dimensionless and $\delta_k(t)$ scales like a length, which we interpret, as before, as the distance between the vector fields u and u_k.

The papers [3,13,14] provide reviews of DiPerna and Lions's theory and of Crippa and De Lellis's contribution.

Obviously, the above result confirms that the earlier principle remains valid: the velocity gradient provides control over the logarithmic relative distance between particle trajectories in the Sobolev case too. Moreover, the rate of convergence of the trajectory ϕ_k to the trajectory ϕ is at least of order $\delta_k(t)$ if the latter is tending to zero.

In this weaker setting, the control of the logarithmic distance ceases to hold uniformly in space. Nevertheless, Crippa and De Lellis are able to deduce local Lipschitz bounds for the generalized flow. (See also [4] for earlier similar results in this direction.) Moreover, uniqueness can be obtained in a similar way to that outlined above in the case of Lipschitz vector fields.

The quantitative theory of Crippa and De Lellis fails to cover the full range of vector fields considered earlier by DiPerna and Lions [20] and Ambrosio [2]. Instead, Crippa and De Lellis had to restrict the setting to Sobolev regular vector fields $u \in L^1((0,T);W^{1,p}(\Omega))$ with $p > 1$. The reason for this is technical: they cleverly exploited standard tools from harmonic analysis (maximal functions) whose strong properties just cease to hold if $p = 1$. Stability estimates in the case $p = 1$ (and also the BV case) are still open. On the positive side, in [25], Jabin manages to extend estimate (2.6) to the $W^{1,1}$ setting modulo an $o(|\log\delta|)$ factor. This estimate is still strong enough to yield uniqueness and stability, but without rates. A direct proof of uniqueness in the BV setting, i.e. without using the uniqueness of the associated partial differential equations as in [2], was (partially) obtained by Jabin [25] and by Hauray and Le Bris [23]. A further extension to the case where the velocity gradient is given by a singular integral is treated by Bouchut and Crippa [5].

It remains to remark that stability estimates in the $p = 1$ case are closely related to a mixing conjecture by Bressan [8]. Indeed, in [11] Crippa and De Lellis derive the $p > 1$ analogue of this conjecture from an estimate similar to (2.6). See also § 3.5 (or [24,37]) for the corresponding result in the Eulerian setting.

3. Stability estimates for continuity equations

In this section, we shall present stability estimates in the Eulerian framework that are similar to the ODE theory in [11]. That is, instead of tracing single particles in a fluid, we shall study the evolution of macroscopic density functions. Our initial intention here is to work out analogies to the Lagrangian framework. We therefore study the case of approximate vector fields in § 3.2. Like the estimates in (2.4) and (2.6), the result will be quite general, as no relation between the two advecting velocity fields is assumed. In § 3.3, we study convergence rates for the zero-diffusivity limit. Section 3.4 is devoted to the convergence order of the numerical upwind scheme. We conclude in § 3.5 with an estimate on mixing rates.

We start by introducing some notation.

3.1. Kantorovich–Rubinstein distance

In order to transfer the Lagrangian stability estimate (2.6) to the Eulerian specification we need some preliminaries. The quantity that will replace Crippa and De Lellis's logarithmic trajectory distance is a Kantorovich–Rubinstein distance with logarithmic cost function taken from the theory of optimal transportation and given by

$$\mathcal{D}_\delta(\rho_1,\rho_2) = \inf_{\pi\in\Pi(\rho_1,\rho_2)} ⨍⨍ \log\left(\frac{|x-y|}{\delta}+1\right)\mathrm{d}\pi(x,y). \tag{3.1}$$

Functionals of this type were initially introduced to model minimal costs for transporting mass from one configuration to the other. For two non-negative distributions ρ_1 and ρ_2, the set $\Pi(\rho_1,\rho_2)$ consists of all transport plans π that realize this transport, i.e.

$$\pi[A\times\Omega] = \int_A \rho_1\,\mathrm{d}x,\quad \pi[\Omega\times A] = \int_A \rho_2\,\mathrm{d}x$$

for any measurable set A. The integrand in (3.1) is the so-called cost function that determines the price for the transport between two points.[1] We refer the interested reader to Villani's monograph [44] for a comprehensive introduction to this topic.

In order to compare this Kantorovich–Rubinstein distance with the trajectory distance considered by Crippa and De Lellis, we note that, in the case where ρ_1 and ρ_2 can be written as push-forwards of the same configuration, which is, for instance, the case if ρ_1 and ρ_2 are transported by different flow fields ϕ_1 and ϕ_2 while having the same initial configuration $\bar\rho$ (cf. (1.3)), $\mathrm{d}\pi = (\phi_1\otimes\phi_2)_\#\delta_{x=y}\otimes\mathrm{d}\bar\rho$ defines an admissible transport plan in $\Pi(\rho_1,\rho_2) = \Pi((\phi_1)_\#\bar\rho,(\phi_2)_\#\bar\rho)$. In particular,

$$\mathcal{D}_\delta(\rho_1,\rho_2) \leqslant ⨍ \log\left(\frac{|\phi_1(x)-\phi_2(x)|}{\delta}+1\right)\bar\rho(x)\,\mathrm{d}x, \tag{3.2}$$

which means that the Kantorovich–Rubinstein distance $\mathcal{D}_\delta(\rho_1,\rho_2)$ is controlled by a weighted variant of Crippa and De Lellis's logarithmic trajectory distance.

Let us now review some of the properties that make Kantorovich–Rubinstein distances convenient in the study of stability estimates for continuity equations. In fact, what is important in our theory is that the quantity $\mathcal{D}_\delta(\rho_1,\rho_2)$ constitutes a mathematical *distance* on the space of configurations with the same total mass [44, theorem 7.3], and it metrizes weak convergence [44, theorem 7.12]. That is,

$$\mathcal{D}_\delta(\rho_k,\rho)\to 0 \quad\Longleftrightarrow\quad \rho_k\to 0 \text{ weakly}. \tag{3.3}$$

[1] Concave cost functions are indeed natural in economics applications as they allow the *economy of scale* to be incorporated into the mathematical model.

If there exists a sequence of δ_k decaying to zero as $k \to \infty$ and such that $\mathcal{D}_{\delta_k}(\rho_k, \rho)$ is uniformly bounded, (3.3) thus yields that ρ_k converges weakly to ρ with a rate not higher than δ_k.

A crucial notion on which the stability analysis of [38] is based is the dual formulation in the Kantorovich–Rubinstein theorem:

$$\mathcal{D}_\delta(\rho_1, \rho_2) = \sup_\zeta \left\{ \fint \zeta(\rho_1 - \rho_2)\,\mathrm{d}x \colon |\zeta(x) - \zeta(y)| \leqslant \log\left(\frac{|x-y|}{\delta} + 1 \right) \right\}$$

(cf. [44, theorem 1.14]). One of its immediate consequences is that $\mathcal{D}_\delta(\rho_1, \rho_2)$ is a transshipment cost that only sees the difference between ρ_1 and ρ_2 ('shared mass stays in place'). We can thus write

$$\mathcal{D}_\delta(\rho_1, \rho_2) = \mathcal{D}_\delta(\rho_1 - \rho_2) \quad \text{or} \quad \mathcal{D}_\delta(\rho) = \mathcal{D}_\delta(\rho^+, \rho^-)$$

where ρ^+ and ρ^- denote, respectively, the positive and negative parts of ρ. It follows that Kantorovich–Rubinstein distances can be considered as distances between any two *not necessarily non-negative* configurations with same the average.

In [38], Seis computed the rate of change of the Kantorovich–Rubinstein distance under the continuity equation with source

$$\partial_t \rho + \nabla \cdot (u\rho) = \nabla \cdot \sigma.$$

By extending some of the earlier techniques developed in [7,11,34,37] in the Lagrangian setting, he found that

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{D}_\delta(\rho) \right| \lesssim_{\Lambda,\bar\rho} \|\nabla u\|_{L^p} + \frac{1}{\delta} \|\sigma\|_{L^1} \tag{3.4}$$

(cf. [38, proposition 1]) if ρ has zero mean.

3.2. Approximating the vector field

We now consider the situation from (2.6). We thus let u and u_k be two vector fields in $L^1(W^{1,p})$ (with $p > 1$) satisfying the compressibility condition (2.5), and we denote by ρ and ρ_k the corresponding solutions to the continuity equation (1.1), starting with the same initial datum $\bar\rho$ in L^q with $1/p + 1/q = 1$. As an immediate consequence of (3.4) and (1.5), we obtain our first result.

Theorem 3.1 (Seis [38]). *If $\delta_k(t)$ denotes the distance between the vector fields u and u_k given by*

$$\delta_k(t) = \int_0^t \|u - u_k\|_{L^p}\,\mathrm{d}t,$$

then

$$\sup_{[0,T]} \mathcal{D}_{\delta_k}(\rho, \rho_k) \lesssim_{\Lambda,\bar\rho} \int_0^T \|\nabla u\|_{L^p}\,\mathrm{d}t + 1 \tag{3.5}$$

holds.

Note that there is similarity to the control principle we found earlier in the ODE case: the velocity gradient controls the logarithmic relative distance between two configurations.

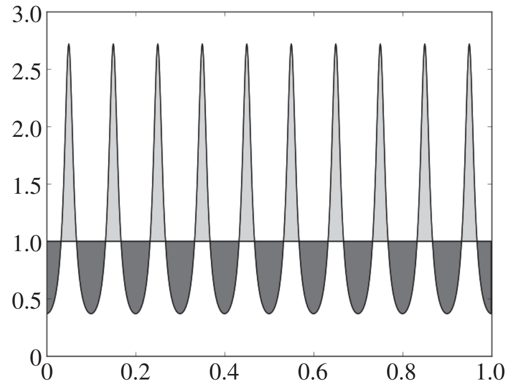

Figure 1. The figure shows the oscillating density ρ_{10} at time $t = 1$. The corresponding Kantorovich–Rubinstein distance $\mathcal{D}_{\delta_{10}}(\rho_{10}, \rho)$ measures the transport between $(\rho_{10} - \rho)^+$ (light grey region) and $(\rho_{10} - \rho)^-$ (dark grey region).

Regarding the fact that Kantorovich–Rubinstein distances metrize weak convergence, in the situation where $\delta_k \to 0$, (3.5) now shows that

$$\rho_k \to \rho \text{ weakly with rate not larger than } \delta_k.$$

This estimate is sharp, as can be seen from the following example, suggested by De Lellis *et al*. [15].

EXAMPLE 3.2. Consider the oscillating vector field $u_k(x) = \sin(2\pi kx)/2\pi k$ on the interval $\Omega = [0, 1]$. Solving the continuity equation with the initial datum $\bar{\rho} = 1$ yields the oscillating solution

$$\rho_k(t, x) = \frac{1 + \tan^2(\pi kx)}{\mathrm{e}^t + \mathrm{e}^{-t}\tan^2(\pi kx)}$$

(see figure 1). Because u_k converges strongly to zero as $k \to \infty$, it is clear that the limiting problem is stationary, i.e. $\rho \equiv 1$.

In view of the oscillatory behaviour of ρ_k, the convergence to $\rho \equiv 1$ is merely weak:

$$\|\rho - \rho_k\|_{L^1(L^1)} \not\to 0.$$

In order to quantify the rate of weak convergence, we note that

$$\delta_k(t) = \int_0^t \|u - u_k\|_{L^p}\,\mathrm{d}t = \frac{t}{2\pi k}\left(\int_0^1 |\sin(2\pi kx)|^p\,\mathrm{d}x\right)^{1/p} \sim \frac{t}{k}$$

because $u = 0$. By the periodicity and symmetry of the problem, we furthermore compute

$$\mathcal{D}_{\delta_k}(\rho, \rho_k) = k\mathcal{D}_{\delta_k}(\rho|_{[0,1/k]}, \rho_k|_{[0,1/k]}) = \mathcal{D}_{\delta_1}(\rho, \rho_1) \sim_t 1,$$

where we have rescaled length in the last identity.

This example shows that theorem 3.1 is sharp in two respects: strong convergence of ρ_k to ρ does not generally hold; the result captures the correct rate of convergence.

Even though the result in theorem 3.1 already appeared in [38], a weaker stability estimate was deduced in [38] in order to replace the unwieldy Kantorovich–Rubinstein distance by a standard negative Sobolev norm. In fact, Seis proved that

$$\|\rho - \rho_k\|_{W^{-1,1}} \lesssim_{\Lambda,\bar\rho,u} \frac{1}{|\log \delta_k(t)|}$$

(see [38, theorem 2]). The new formulation in theorem 3.1 has the advantage that it is sharp and naturally extends the analogous estimates in the Lagrangian setting (2.6).

3.3. The zero-diffusivity limit

In this subsection, we expand the model (1.1) by a second parallel transport mechanism in addition to advection, namely diffusion. Advection–diffusion models are ubiquitous in thermodynamics, fluid dynamics and engineering, e.g. in the context of thermal convection [41], spinodal decomposition [40] or mixing [42]. While convection enhances the efficient transport of particles or fluid parcels over large distances and tends to create sharp gradients in the density (or temperature) distribution, diffusion compensates density (or temperature) differences locally.

In the following, we thus consider the Cauchy problem for the advection–diffusion equation:

$$\left.\begin{aligned} \partial_t \rho_\kappa + \nabla\cdot(u\rho_\kappa) &= \kappa\Delta\rho_\kappa && \text{in } (0,\infty)\times\Omega, \\ \rho_\kappa(0,\cdot) &= \bar\rho && \text{in } \Omega, \end{aligned}\right\} \tag{3.6}$$

where κ is the (positive) diffusivity constant. Equipping the equation with the no-flux condition $\nabla\rho_\kappa\cdot\nu = 0$ on the boundary of Ω implies that the evolution is still mass conserving:

$$\int \rho_\kappa(t,x)\,\mathrm{d}x = \int \bar\rho\,\mathrm{d}x. \tag{3.7}$$

We assume furthermore that $\bar\rho$ is non-negative, and so ρ_κ is as a consequence of the maximum principle for (3.6). Note that (1.5) remains valid for ρ_κ, which can easily be seen by testing (3.6) with ρ_κ^{q-1}.

We are interested in the vanishing diffusivity limit $\kappa\to 0$. In order to quantify the rate of convergence of solutions of (3.6) towards solutions of the purely advective model (1.1), we shall make use of a standard decay estimate from relaxation theory for the diffusion (or heat) equation. A common way to identify the equilibration rate in the diffusive model is by studying the decay behaviour of the entropy:

$$H(\rho) = \fint \rho\log\rho\,\mathrm{d}x.$$

We compute the rate of change of entropy under the evolution equation (3.6) using multiple integrations by parts:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}H(\rho_\kappa) &= \kappa\fint \Delta\rho_\kappa\log\rho_\kappa\,\mathrm{d}x - \fint \nabla\cdot(u\rho_\kappa)\log\rho_\kappa\,\mathrm{d}x \\ &= -\kappa\fint \frac{|\nabla\rho_\kappa|^2}{\rho_\kappa}\,\mathrm{d}x - \fint(\nabla\cdot u)\rho_\kappa\,\mathrm{d}x, \end{aligned}$$

where in the first equality we have used the fact that the evolution is mass conserving (3.7). Moreover, since ρ_κ is a non-negative function, integration in time yields

$$H(\rho_\kappa(t,\cdot)) + \kappa \int_0^t \fint \frac{|\nabla \rho_\kappa|^2}{\rho_\kappa}\,\mathrm{d}x\,\mathrm{d}t \leqslant H(\bar\rho) + \left(\int_0^t \|(\nabla\cdot u)^-\|_{L^\infty}\,\mathrm{d}t\right)\|\rho_\kappa\|_{L^\infty(L^1)}$$

for any $t \in (0,T)$. Then, if the initial density has finite entropy, by the Hölder inequality and mass conservation (3.7),

$$\int_0^t \fint |\nabla\rho_\kappa|\,\mathrm{d}x\,\mathrm{d}t \leqslant \int_0^t \left(\fint \rho_\kappa\,\mathrm{d}x \fint \frac{|\nabla\rho_\kappa|^2}{\rho_\kappa}\,\mathrm{d}x\right)^{1/2}\mathrm{d}t \lesssim_{\bar\rho,\Lambda} \sqrt{\frac{t}{\kappa}}.$$

Using the theory developed in [38], cf. (3.4), and the a priori estimate (1.5), it can now be shown that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{D}_\delta(\rho,\rho_\kappa) \lesssim_{\Lambda,\bar\rho} \|\nabla u\|_{L^p} + \frac{\kappa}{\delta}\|\nabla\rho_\kappa\|_{L^1}.$$

Integration in time and a combination of the previous two estimates then yield the following result.

THEOREM 3.3. *Let $\delta_\kappa(t)$ be the diffusion distance per time t, i.e. $\delta_\kappa(t) = \sqrt{t\kappa}$. Then*

$$\sup_{[0,T]} \mathcal{D}_{\delta_\kappa}(\rho,\rho_\kappa) \lesssim \int_0^T \|\nabla u\|_{L^p} + 1.$$

In other words, the diffusive approximation converges weakly to the unique solution of the continuity equation with a rate not larger than $\sqrt{t\kappa}$. The latter is approximately equal to the distance a particle can travel by diffusion in time t. A qualitative convergence result was previously established for this case in [20] but, to the best of our knowledge, this is the first time that a convergence rate for the zero-diffusivity limit has been obtained.

In fact, a simple calculation shows that theorem 3.3 is optimal.

EXAMPLE 3.4. Consider the advection–diffusion equation on the real line with constant velocity $u \equiv U > 0$ and initial datum

$$\bar\rho(x) = \begin{cases} x^{-(1-\varepsilon)} & \text{for } x \in (0,1),\\ 0 & \text{otherwise,}\end{cases}$$

where $\varepsilon \in (0,1)$. The corresponding solution is then of the form

$$\rho_\kappa(t,x) = \frac{1}{\sqrt{4\pi\kappa t}}\int_0^1 \exp\left(-\frac{|x-tU-y|^2}{4\kappa t}\right)\frac{1}{y^{1-\varepsilon}}\,\mathrm{d}y,$$

and strongly converges to the solution $\rho(t,x) = \bar\rho(x-tU)$ of the continuity equation. A short calculation then shows that $\|\rho_\kappa - \rho\|_{L^1((0,1)\times\mathbb{R})} \gtrsim \sqrt{\kappa}^{\varepsilon}$, and thus, for any $\varepsilon \in (0,1)$, there is an initial configuration such that

$$\sqrt{\kappa}^{-\varepsilon}\|\rho_\kappa - \rho\|_{L^1(L^1)} \not\to 0.$$

This estimate entails that, for general rough initial data, universal (power-law) convergence rates cannot exist in a strong norm. It is thus natural to investigate the order of weak convergence. For the 1-Wasserstein distance

$$W_1(\rho_1,\rho_2) = \inf_{\pi\in\Pi(\rho_1,\rho_2)} ⨍⨍ |x-y|\,\mathrm{d}\pi(x,y),$$

which controls the logarithmic Kantorovich–Rubinstein distance $\mathcal{D}_\delta(\rho_1,\rho_2)$ with the factor δ^{-1}, we compute the lower bound

$$W_1(\rho,\rho_\kappa) \gtrsim \sqrt{\kappa}^{1-\varepsilon}.$$

Since Wasserstein distances metrize weak convergence [44, theorem 7.12], this estimate shows that ρ_κ converges weakly to ρ with rate at most $\sqrt{\kappa}^{1-\varepsilon}$. Even though the measures of weak convergence differ, this lower bound on the rate of weak convergence (almost) matches the upper bound found in theorem 3.3. (For details on the above estimates, we refer the reader to the very similar computations in [36, § 7].)

3.4. Convergence rates for the upwind scheme

The upwind scheme is a numerical scheme for approximating solutions to the continuity equation. The scheme is a finite volume scheme, which means that the domain is decomposed into control volumes (or cells) of small diameter and the evolving density is approximated by averages over each control volume.

To be more specific, we consider a domain Ω that can be written as a finite union of rectangular boxes. We decompose Ω into a family of rectangular cells with disjoint interiors, $\Omega = \bigcup_{K\in\mathcal{T}} K$, where $\mathcal{T}$ is the tessellation and K is a translation of the cube $[0,h_1]\times\cdots\times[0,h_d]$. The size h of the tessellation is the maximal edge length, i.e.

$$h = \max_{i=1}^{d} h_i.$$

We suppose that the tessellation is regular in the sense that $h_i \sim h$ for all i. For two neighbouring cells $K\sim L$, we denote by $K|L$ the joint boundary. The normal vector on $K|L$ pointing from K to L is denoted by ν_{KL}.

We choose a fixed time step size δt so that the nth time step reads $t^n = n\delta t$. To guarantee the stability of the explicit scheme, we impose the following Courant–Friedrichs–Lewy condition on the time step size:

$$\forall n, \quad \int_{t^n}^{t^{n+1}} \|u\|_{L^\infty}\,\mathrm{d}t \leqslant h.$$

We thus assume in this subsection that $u\in L^1(L^\infty)$.

To approximate the transport term, we consider the net flow from K to L defined by

$$u_{KL}^n = ⨍_{t^n}^{t^{n+1}} ⨍_{K|L} u\cdot\nu_{KL}\,\mathrm{d}\mathcal{H}^{d-1}\,\mathrm{d}t.$$

We remark that these quantities are well defined owing to the trace estimate for Sobolev functions. Furthermore, the initial configuration of the scheme is the volume

https://doi.org/10.1017/S0308210518000173 Published online by Cambridge University Press

average over each $K \in \mathcal{T}$, i.e.

$$\rho_K^0 = \fint_K \bar\rho \, \mathrm{d}x.$$

We are now in a position to define the explicit upwind finite volume scheme for the continuity equation (1.1):

$$\rho_K^{n+1} = \rho_K^n + \frac{\mathrm{d}t}{h} \sum_{L \sim K} (u_{LK}^{n+} \rho_L^n - u_{KL}^{n+} \rho_K^n), \tag{3.8}$$

where $u_{KL}^{n+} = (u_{KL}^n)^+$. The approximate solution is given by

$$\rho_h(t,x) = \rho_K^n \quad \text{if } (t,x) \in [t^n, t^{n+1}) \times K, \ K \in \mathcal{T}.$$

See [22, 28] for properties and references. Under the DiPerna–Lions setting, convergence of the scheme, i.e. $\rho_h \to \rho$ as $h \to 0$, is proved in [6, 36, 45].

Even though the numerical scheme is formally first order, one observes a breakdown in the convergence rate to order $\frac{1}{2}$ in the case of non-smooth initial data. $\sqrt{h}$-rates were numerically observed in [6, 36] for the DiPerna–Lions setting considered here. The reason for this lack of convergence is the occurrence of numerical diffusion that smooths out sharp interfaces. Such irregularities, however, are simply transported in the continuous model. In a certain sense, approximate solutions show a behaviour similar to those of the advection–diffusion equation (3.6), where $\kappa \sim h$. It is this similarity that determines the $\sqrt{h}$-rate of convergence (cf. theorem 3.3). The effect of numerical diffusion is illustrated in [36, § 2.4].

In the case of regular (i.e. at least spatially Lipschitz continuous) vector fields, this breakdown in the order of convergence has long been known. The first rigorous results on optimal convergence rates date back to the 1970s (see, for example, [9, 16, 17, 19, 27, 32, 33, 35, 43]). To the best of our knowledge, the only available result in the DiPerna–Lions setting is very recent: Schlichting and Seis [36] established an upper bound on the rate of weak convergence that captures the optimal order.

THEOREM 3.5 (Schlichting and Seis [36]). *Let $\delta_h(t)$ be the numerical diffusion distance per time t, i.e.*

$$\delta_h(t) = \sqrt{h \int_0^t \|u\|_{L^\infty} \, \mathrm{d}t}.$$

Then

$$\sup_{[0,T]} \mathcal{D}_{\delta_h}(\rho, \rho_h) \lesssim_{\Lambda, \bar\rho} \int_0^T \|\nabla u\|_{L^p} \, \mathrm{d}t + 1.$$

The work of Schlichting and Seis [36] builds not only on the quantitative theory from [38] but on the probabilistic interpretation of the upwind scheme suggested by Delarue *et al*. [16, 17]. In fact, in [36], (3.8) is interpreted as a Markov chain, which comes as a time-discretized version of the stochastic differential equation

$$\mathrm{d}\psi_t = u(t, \psi_t) \, \mathrm{d}t + \sqrt{2h} \, \mathrm{d}W_t,$$

with a noise term depending on the details of the mesh. In a certain sense, the above equation is the Lagrangian analogue of the advection–diffusion equation (3.6). It turns out that the noise term determines the $\sqrt{h}$-rate of convergence.

In fact, the result of [36] is optimal.

EXAMPLE 3.6 (Schlichting and Seis [36]). Consider the continuity equation on the real line with constant velocity $u \equiv U > 0$ and the same singular initial datum as in example 3.4. If the control volumes are taken to be $K = [k, k+1)h$ for $k \in \mathbb{Z}$ and $\delta t U = \frac{1}{2}h$ is the time step size, the upwind scheme reduces to $\rho_k^{n+1} = \frac{1}{2}(\rho_k^n + \rho_{k-1}^n)$, and thus, by iteration

$$\rho_k^n = \frac{1}{2^n} \sum_{m=0}^{n} \binom{n}{m} \rho_{k-m}^0.$$

As a consequence of the numerical diffusion, one cannot expect strong (power-law) convergence rates to hold for rough initial data. Indeed,

$$\sqrt{h}^{-\varepsilon} \|\rho_h - \rho\|_{L^1(L^1)} \not\to 0.$$

On the other hand, in view of the lower bound

$$W_1(\rho, \rho_h) \gtrsim \sqrt{h}^{1-\varepsilon},$$

where W_1 denotes the 1-Wasserstein distance introduced in example 3.4, we find that, for this particular example, the rate of weak convergence is at most of order $\frac{1}{2}(1-\varepsilon)$. Even though the measures of weak convergence differ, this lower bound on $W_1(\rho, \rho_h)$ and the upper bound on $D_{\delta_h}(\rho, \rho_h)$ in theorem 3.5 indicate that the optimal rate of weak convergence is indeed of order $\frac{1}{2}$.

3.5. Mixing by stirring

In recent years, mixing by stirring attracted much interest in both the applied mathematics and engineering communities. Mixing refers to the homogenization of an inhomogeneous substance by being stirred by an agent. The major goals are the quantification of mixing rates and the design of mixing strategies. In order to optimize mixing strategies, absolute lower bounds on the mixing rate are indispensable. In this subsection, we present a lower bound on mixing by the stirring of incompressible viscous fluids obtained in [37]. A nice review on the mathematical side of mixing was written by Thiffeault [42].

A natural constraint in the experimental mixing set-up is the amount of mechanical work the engineer is willing to do in order to overcome viscous friction to maintain stirring. Mathematically, this amounts to limiting the budget of the viscous dissipation rate (or enstrophy) given by $\|\nabla u\|_{L^2}$. In the following, we shall slightly generalize this constraint by assuming that $u \in L^1(W^{1,p})$ for some $p > 1$ as in the previous part of this paper.

While our intuition is strong about whether a substance is well mixed or not, the choice of a measure that quantifies the degree of mixedness depends on the mathematical communities. Homogeneous negative Sobolev norms, in particular the $\dot{H}^{-1/2}$ norm [30,31] and the $\dot{H}^{-1}$ norm [21,29,39] are favoured by fluid dynamicists. These norms measure oscillations: the greater the length scales, the larger the norms.

In [37] Seis introduces a new mixing measure: a variant of the Kantorovich–Rubinstein distance introduced earlier in the present paper. We accordingly consider

$$M(\rho) = \inf_{\pi \in \Pi(\rho^+,\rho^-)} \exp\left(\fint\fint \log|x-y| \,\mathrm{d}\pi(x,y) \right).$$

Note that $M(\rho) = \lim_{\delta\to 0} \exp(\mathcal{D}_\delta(\rho) + (\log\delta)\|\rho\|_{L^1})$. In the case of a two-phase mixture, modelled by $\rho \in \{\pm 1\}$, this distance formally scales as a length, so $M(\rho)$ heuristically agrees with the average size of the unmixed regions.

The mixing process can be modelled by the continuity equation (1.1), which turns into the transport equation

$$\partial_t \rho + u \cdot \nabla\rho = 0 \tag{3.9}$$

under the assumption that the fluid is incompressible (i.e. $\nabla \cdot u = 0$), which we shall make for convenience. For simplicity, we restrict our attention to two-phase mixtures with equal volume fraction, so that

$$|\{x \in \Omega\colon \rho(t,x) = 1\}| = |\{x \in \Omega\colon \rho(t,x) = -1\}|$$

for any $t > 0$, or, equivalently, $\int \rho \,\mathrm{d}x = 0$.

Seis [37] derives a lower bound on mixing rates in incompressible viscous fluids, building on an estimate similar to (3.4).

Theorem 3.7 (Seis [37]). *For any $T \geqslant 0$, it holds that*

$$M(\rho(T,\cdot)) \geqslant M(\bar\rho) \exp\left(-\frac{1}{C} \int_0^T \|\nabla u\|_{L^p} \,\mathrm{d}t \right), \tag{3.10}$$

where C is a constant depending only on p and d.

This estimate shows the impossibility of perfect mixing, i.e. $\rho \to 0$ weakly, in finite time. A similar statement was obtained earlier by Crippa and De Lellis [11] for a certain geometric mixing measure suggested by Bressan [8]. In fact, Bressan conjectures the $p = 1$ analogue of Crippa and De Lellis's estimate.

It is not difficult to deduce a lower bound on the decay rate of the $\dot H^{-1}$ norm from (3.10). Indeed, in [37] it is proved that

$$\frac{1}{|\rho|_{\mathrm{BV}}} \lesssim M(\rho) \leqslant |\rho|_{H^{-1}}, \tag{3.11}$$

where $|\rho|_{\mathrm{BV}}$ and $|\rho|_{H^{-1}}$ denote, respectively, the homogeneous part of the BV norm (and thus $|\rho|_{\mathrm{BV}} = 2|\partial\{\rho = 1\}|/|\Omega|$) and the homogeneous part of the H^{-1} norm. The first inequality in (3.11) is an interpolation inequality, whereas the second follows immediately via Jensen's inequality and the Kantorovich–Rubinstein theorem [44, theorem 1.14]. Plugging (3.11) into theorem 3.7 yields

$$|\rho(T,\cdot)|_{H^{-1}} \gtrsim \frac{1}{|\bar\rho|_{\mathrm{BV}}} \exp\left(-\frac{1}{C} \int_0^T \|\nabla u\|_{L^p} \,\mathrm{d}t \right). \tag{3.12}$$

A similar decay estimate for the $\dot H^{-1}$ norm was obtained simultaneously by Iyer *et al*. [24] by using the geometric results from [11].

Estimates (3.10) and (3.12) are sharp. This was proved independently by Yao and Zlatoš [46] and by Alberti *et al*. [1]. In fact, in both works, explicit mixing flows are constructed that saturate the lower bounds from [24, 37]. Numerical evidence for the optimality of this mixing rate was given in [29].

There is a close relation between theorem 3.7 and the lower bound in (2.1); in fact, the upper bound

$$M(\rho(T,\cdot)) \leqslant M(\bar\rho)\exp\left(\frac{1}{C}\int_0^T \|\nabla u\|_{L^p}\,\mathrm{d}t\right)$$

is also valid. Estimate (3.10) can be seen as the Eulerian (and Sobolev) analogue of (2.1), in the sense that in theorem 3.7 we compute the distance between the configuration described by the mixing process and the stationary fully mixed state $\rho = 0$. While (2.1) shows that trajectories cannot converge faster than exponentially in time, the Eulerian analogue shows that different density configurations cannot converge faster than exponentially in time. This observation also underlines the link between mixing and the question of uniqueness for the partial differential equation (1.1) (or (3.9)): a system is perfectly mixing in finite time precisely if solutions to (1.1) are generally not unique. Note that, in the case of finite-time mixing, upon reversing time, one gains non-trivial solutions to (1.1) with initial datum zero. An explicit construction of such an unmixing solution is due to Depauw [18].

It remains to note that upper bounds on the rates of unmixing (or coarsening) in viscous fluids were obtained in [7, 34]; the analysis in these papers combines (3.10) and the lower bound of (3.11) with the Kohn–Otto upper bound method [26].

Acknowledgements

The author thanks André Schlichting for fruitful discussions and for suggesting the entropy approach in § 3.3.

References

1 G. Alberti, G. Crippa and A. L. Mazzucato. Exponential self-similar mixing and loss of regularity for continuity equations. *C. R. Math.* **352** (2014), 901–906.

2 L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.* **158** (2004), 227–260.

3 L. Ambrosio and G. Crippa. Continuity equations and ODE flows with non-smooth velocity. *Proc. R. Soc. Edinb.* A **144** (2014), 1191–1244.

4 L. Ambrosio, M. Lecumberry and S. Maniglia. Lipschitz regularity and approximate differentiability of the DiPerna–Lions flow. *Rend. Sem. Mat. Univ. Padova* **114** (2005), 29–50.

5 F. Bouchut and G. Crippa. Lagrangian flows for vector fields with gradient given by a singular integral. *J. Hyperbol. Diff. Eqns* **10** (2013), 235–282.

6 F. Boyer. Analysis of the upwind finite volume method for general initial- and boundary-value transport problems. *IMA J. Numer. Analysis* **32** (2012), 1404–1439.

7 Y. Brenier, F. Otto and C. Seis. Upper bounds on coarsening rates in demixing binary viscous liquids. *SIAM J. Math. Analysis* **43** (2011), 114–134.

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

9 B. Cockburn, B. Dong, J. Guzmán and J. Qian. Optimal convergence of the original DG method on special meshes for variable transport velocity. *SIAM J. Numer. Analysis* **48** (2010), 133–146.

10 M. Colombo, G. Crippa and S. Spirito. Renormalized solutions to the continuity equation with an integrable damping term. *Calc. Var. PDEs* **54** (2015), 1831–1845.

11 G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna–Lions flow. *J. Reine Angew. Math.* **616** (2008), 15–46.

12 G. Crippa, C. Nobili and C. Seis. Eulerian and Lagrangian solutions to the continuity and Euler equations with L^1 vorticity. *SIAM J. Math. Analysis* **49** (2017), 3973–3998.

13 C. De Lellis. ODEs with Sobolev coefficients: the Eulerian and the Lagrangian approach. *Discrete Contin. Dyn. Syst.* S **1** (2008), 405–426.

14 C. De Lellis. *Ordinary differential equations with rough coefficients and the renormalization theorem of Ambrosio*. Astérisque, vol. 317, pp. 175–203 (Montrouge: Société Mathématique de France, 2008).

15 C. De Lellis, P. Gwiazda and A. Świerczewska-Gwiazda. Transport equations with integral terms: existence, uniqueness and stability. *Calc. Var. PDEs* **55** (2016), 128. (Available at https://doi.org/10.1007/s00526-016-1049-9.)

16 F. Delarue and F. Lagoutière. Probabilistic analysis of the upwind scheme for transport equations. *Arch. Ration. Mech. Analysis* **199** (2011), 229–268.

17 F. Delarue, F. Lagoutière and N. Vauchelet. Convergence order of upwind type schemes for transport equations with discontinuous coefficients. *J. Math. Pures Appl.* **108** (2017), 918–951.

18 N. Depauw. Non-unicité du transport par un champ de vecteurs presque BV. In *Séminaire: Équations aux Dérivées Partielles, 2002–2003*. Séminaire Laurent Schwartz ÉDP et Applications, vol. 19 (Palaiseau: Editions École Polytechique, 2003).

19 B. Despres. Lax theorem and finite volume schemes. *Math. Comput.* **73** (2004), 1203–1234.

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

21 C. R. Doering and J.-L. Thiffeault. Multiscale mixing efficiencies for steady sources. *Phys. Rev.* E **74** (2006), 025301.

22 R. Eymard, T. Gallouët and R. Herbin. Finite volume methods. In *Handbook of numerical analysis*, vol. VII, pp. 713–1020 (Amsterdam: North-Holland, 2000).

23 M. Hauray and C. Le Bris. A new proof of the uniqueness of the flow for ordinary differential equations with BV vector fields. *Annali Mat. Pura Appl.* **190** (2011), 91–103.

24 G. Iyer, A. Kiselev and X. Xu. Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows. *Nonlinearity* **27** (2014), 973–985.

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

26 R. V. Kohn and F. Otto. Upper bounds on coarsening rates. *Commun. Math. Phys.* **229** (2002), 375–395.

27 N. N. Kuznecov. The accuracy of certain approximate methods for the computation of weak solutions of a first order quasilinear equation. *Zh. Vychisl. Matem. Mat. Fiz.* **16** (1976), 1489–1502.

28 R. J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics (Cambridge University Press, 2002).

29 Z. Lin, J.-L. Thiffeault and C. R. Doering. Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.* **675** (2011), 465–476.

30 G. Mathew, I. Mezić and L. Petzold. A multiscale measure for mixing. *Physica* D **211** (2005), 23–46.

31 G. Mathew, I. Mezić, S. Grivopoulos, U. Vaidya and L. Petzold. Optimal control of mixing in Stokes fluid flows. *J. Fluid Mech.* **580** (2007), 261–281.

32 B. Merlet. L^∞- and L^2-error estimates for a finite volume approximation of linear advection. *SIAM J. Numer. Analysis* **46** (2007), 124–150.

33 B. Merlet and J. Vovelle. Error estimate for finite volume scheme. *Numer. Math.* **106** (2007), 129–155.

34 F. Otto, C. Seis and D. Slepčev. Crossover of the coarsening rates in demixing of binary viscous liquids. *Commun. Math. Sci.* **11** (2013), 441–464.

35 T. E. Peterson. A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation. *SIAM J. Numer. Analysis* **28** (1991), 133–140.

36 A. Schlichting and C. Seis. Convergence rates for upwind schemes with rough coefficients. *SIAM J. Numer. Analysis* **55** (2017), 812–840.

37 C. Seis. Maximal mixing by incompressible fluid flows. *Nonlinearity* **26** (2013), 3279–3289.

38 C. Seis. A quantitative theory for the continuity equation. *Annales Inst. H. Poincaré Analyse Non Linéaire* **34** (2017), 1837–1850.

39 T. A. Shaw, J.-L. Thiffeault and C. R. Doering. Stirring up trouble: multi-scale mixing measures for steady scalar sources. *Physica* D **231** (2007), 143–164.

40 E. D. Siggia. Late stages of spinodal decomposition in binary mixtures. *Phys. Rev.* A **20** (1979), 595–605.

41 E. D. Siggia. High Rayleigh number convection. *A. Rev. Fluid Mech.* **26** (1994), 137–168.

42 J.-L. Thiffeault. Using multiscale norms to quantify mixing and transport. *Nonlinearity* **25** (2012), R1–R44.

43 J.-P. Vila and P. Villedieu. Convergence of an explicit finite volume scheme for first order symmetric systems. *Numer. Math.* **94** (2003), 573–602.

44 C. Villani. *Topics in optimal transportation*. Graduate Studies in Mathematics, vol. 58 (Providence, RI: American Mathematical Society, 2003).

45 N. J. Walkington. Convergence of the discontinuous Galerkin method for discontinuous solutions. *SIAM J. Numer. Analysis* **42** (2005), 1801–1817.

46 Y. Yao and A. Zlatoš. Mixing and un-mixing by incompressible flows. *J. Eur. Math. Soc.* **19** (2017), 1911–1948.