SCUOLA
NORMALE
SUPERIORE

Classe di Scienze

Corso di perfezionamento in Metodi Computazionali e Modelli Matematici per le Scienze e la Finanza

XXXVII ciclo

Continuity and Transport Equations: Theory, Numerics & Modeling

Settore Scientifico Disciplinare **MAT/05**

Candidato
Dr. Tommaso Cortopassi

Relatori

Prof. Giovanni Alberti
Prof. Antonio De Simone

Supervisor

Prof. Franco Flandoli

Anno accademico 2024–2025

Contents

Introduction

It would be easy to say that the main topic of this thesis is the study of the *continuity equation* and of the *transport equation*. It would not be an incorrect statement. However, we think that the main purpose of this work, conveyed through the analysis of these PDEs, is to show how the Eulerian and the Lagrangian points of view are complementary: results which seem out of reach from one perspective can often be established by working with the other.

As we mentioned, the specific PDEs we are going to study are the continuity equation

$$\begin{cases} \partial_t u + \operatorname{div}_x(bu) = 0 & \text{in } \mathbb{T}^d \text{ with } t > 0 \\ u(0,\cdot) = u_0 & \text{for t=0} \end{cases} \tag{0.0.1}$$

and the transport equation

$$\begin{cases} \partial_t \rho + b \cdot \nabla_x \rho = 0 & \text{in } \mathbb{T}^d \text{ with } t > 0 \\ \rho(0,\cdot) = \rho_0 & \text{for t=0} \end{cases}, \tag{0.0.2}$$

where b is a given velocity field and u_0, ρ_0 are initial data. If the velocity field b is "regular enough" the well-posedness theory for both equations is classical and linked to the Cauchy problem

$$\begin{cases} \frac{d}{dt}\gamma(t) = b(t,\gamma(t)) \\ \gamma(t_0) = x_0 \end{cases}. \tag{0.0.3}$$

To be more precise, a connection can be established via the *flow* of b, i.e. a map $X : [0,T) \times \mathbb{T}^d \to \mathbb{T}^d$ which is a "collection" of integral curves of b, i.e. $X(\cdot, x)$ solves

$$\begin{cases} \frac{d}{dt}X(t,x) = b(t,X(t,x)) \\ X(0,x) = x \end{cases} \tag{0.0.4}$$

for every $x \in \mathbb{T}^d$ [1]. Namely, for every $x \in \mathbb{T}^d$ and assuming that b is regular, the map $X(\cdot,x) : [0,T) \to \mathbb{T}^d$ is the solution of (0.0.3), which exists and is unique by the classical Cauchy theory for ODEs. The connection between the continuity and the transport equations and the flow X of b can be described as follows: the solutions u and ρ can be obtained by transporting the initial data u_0 and ρ_0 along the integral curves $t \mapsto X(t,x)$, a technique known in the literature as *method of characteristics*. Even though this description highlights the similarity between the two equations, and in some aspects they are indeed similar, they are also fundamentally different, as they describe the evolution of inherently distinct types of quantities. We can summarise this difference by saying that the continuity equation describes the evolution of *densities* (or, more generally, measures), while the transport equation describes the evolution of *scalars*. A good example (though highly simplified) to illustrate this distinction is the following. If we introduce a drop of ink into a river, the ink particles will follow the streamlines of the water. Over time, the drop of ink will change shape and eventually form a blob in which different regions have different ink densities. The density of the ink thus changes over time, which is precisely what the continuity equation describes. On the other hand, if we throw some grains of sand into the same river, the grains will still follow the water streamlines just like the ink did, but their density will not change: each grain is discrete and at any given time and position

[1] We immediately inform the reader that, throughout this thesis, we will almost always work in a d-dimensional torus $\mathbb{T}^d$ rather than in $\mathbb{R}^d$. This is only a technical assumption, as we shall explain.

either a grain is present or it is not. In this case the initial distribution is merely rearranged, which is what is described by the transport equation.

Outside of the regular case (i.e. when b is not Lipschitz continuous in space) the situation remained unclear for several years, since non-uniqueness phenomena may occur. A breakthrough was obtained by DiPerna and Lions in their seminal paper [70], where they proved well-posedness of (0.0.4) assuming Sobolev spatial regularity of the velocity field b and a control on the $L^1_t L^\infty_x$ norm of its divergence. Contrary to the strategy followed in the classical case with the method of characteristics, they obtain the well-posedness of (0.0.4) by first proving well-posedness at the PDE level for (0.0.1) and (0.0.2) [2]. The main problem is proving *uniqueness* of solutions and the strategy they use is the so-called *renormalization technique*. We will not go into details here, we just give an idea using the transport equation (0.0.2) as an example, even though the same argument applies also to (0.0.1), with minor modifications. In the regular setting, if we consider the composition of a solution ρ of (0.0.2) with a function $\beta \in C^1(\mathbb{R};\mathbb{R})$, by the chain-rule $\beta(\rho)$ solves

$$\partial_t[\beta(\rho)] + b \cdot \nabla_x[\beta(\rho)] = 0. \tag{0.0.5}$$

Notice that, outside the regular setting, even though the existence of a solution ρ for (0.0.2) can be proved rather easily we have no information on its regularity, so the chain-rule may not apply and $\beta(\rho)$ may not be a solution of (0.0.5). If for every $\beta \in C^1(\mathbb{R};\mathbb{R})$ it holds that $\beta(\rho)$ solves (0.0.5), then we say that ρ is a *renormalized solution*, and it can be proved that renormalized solutions are unique. The key property is the *renormalization property* for the vector field b, which states that *every* weak solution of (0.0.2) is a renormalized solution, and thus unique. Notice how the focus has shifted from the solution ρ to the vector field b: we only need to prove that b satisfies the renormalization property to have, as an easy corollary, that weak solutions of (0.0.2) are unique. This is achieved through a regularisation scheme which shows that vector fields satisfying certain regularity conditions (namely Sobolev regularity in the spatial variables and a boundedness assumption on the divergence) have the renormalization property. With such well-posedness results at the PDE level, corresponding well-posedness results at the ODE level for solutions of (0.0.4) can also be established. In particular, it is important to point out that the notion of uniqueness for solutions of (0.0.4) in this non-regular setting is a weak one: it is not necessarily true that for every $x \in \mathbb{T}^d$ there exists a unique solution $X(\cdot, x)$ of (0.0.3). A general abstract framework in which well-posedness results at the PDE level can be automatically translated into a "weak" well-posedness for ODEs was provided by Ambrosio [7]. In the same work, the author also extended the results in [70] by proving that vector fields with BV spatial regularity (with a boundedness assumption on the divergence) have the renormalization property. In particular, the appropriate notion of flow to be considered as a solution of (0.0.4) in this weaker setting is the so-called *regular Lagrangian flow*. The underlying idea is that a "good" solution X of (0.0.4) is one which does not compress the Lebesgue measure too much; in other words, if we transport bounded initial data via the flow X, the boundedness is preserved. When restricting to this "good" class of flows, uniqueness at the PDE level implies uniqueness at the ODE level.

In more recent years, the topic of numerically approximating the solutions of (0.0.1), (0.0.2) and (0.0.4) has also been investigated. Indeed, the interest for a well-posedness theory in a low-regularity setting was not only motivated by the will of extending the classical theory, but also to study PDEs arising from physics (e.g. fluid mechanics) where the vector fields are often not regular enough to apply the classical theory. However, there is a problem. The theory developed by DiPerna-Lions and Ambrosio is inherently *qualitative*. Namely, if we consider a smooth approximation $b_\varepsilon \xrightarrow{\varepsilon \to 0} b$ of a vector field b for which well-posedness of (0.0.1),(0.0.2) and (0.0.4) holds, the theory of renormalized solutions provides no rate of convergence either for the flow X_ε satisfying

$$\partial_t X_\varepsilon(t,x) = b_\varepsilon(t, X_\varepsilon(t,x))$$

or for the solution ρ_ε of

$$\partial_t \rho_\varepsilon + b_\varepsilon \cdot \nabla_x \rho_\varepsilon = 0$$

[2]In fact, they prove well-posedness for a more general PDE which (0.0.1) and (0.0.2) are special cases of.

to the regular Lagrangian flow X of b and the solution ρ of (0.0.2), respectively. Quantitative stability estimates for regular Lagrangian flows were first obtained by Crippa and De Lellis [58] using completely different techniques. It is worth noticing that, with their arguments, they can prove directly the well-posedness of (0.0.4) without having to rely on the well-posedness at the PDE level. However, for technical reasons, their argument needs stronger assumptions on the vector field. Indeed their proofs rely on three main properties:

1. A pointwise inequality (which holds almost everywhere) for BV functions;
2. the boundedness of the maximal operator from L^p to L^p;
3. a uniform lower bound on $\det(\nabla_x X(t,\cdot))$.

Vector fields satisfying DiPerna-Lions (and even Ambrosio) regularity assumptions satisfy the first and third properties, but the second one requires $p > 1$. For this reason, up to now a quantitative theory (both at the PDE and at the ODE level) is, in general, limited to $L^1_t W^{1,p}_x$ vector fields with $p > 1$. In particular, in [58] it is proved that given vector fields b_1 and b_2 satisfying the necessary assumptions and given their regular Lagrangian flows X_1 and X_2, then

$$\sup_t \|X_1(t,\cdot) - X_2(t,\cdot)\|_{L^1_x} \lesssim |\log(\|b_1 - b_2\|_{L^1_t L^1_x})|^{-1}. \tag{0.0.6}$$

Once again shifting the focus from the Lagrangian/ODE point of view to the Eulerian/PDE one, using the quantitative estimates provided in [58], Seis [144] established corresponding estimates at the PDE level, providing quantitative stability results for solutions of (0.0.1) with respect to the Wasserstein distance.

With quantitative estimates at hand, one can devise numerical approximation schemes with explicit convergence rates for solutions of (0.0.1). This was first accomplished in [142], where the authors analysed an explicit upwind finite-volume scheme. In more recent years, other works have addressed the approximation of solutions of (0.0.1) in non-regular settings, though always adopting an Eulerian perspective on the problem. In [56], the author of this thesis first established a numerical approximation scheme (namely, an explicit Euler method) for approximating the regular Lagrangian flow of a vector field, proving a quantitative estimate analogous to (0.0.6), and then used these Lagrangian estimates to derive quantitative approximation results for solutions of (0.0.1). A generalisation of these results, employing a θ-method instead of an explicit Euler method, is obtained by the author and collaborators in [46].

A very similar connection between the Lagrangian and the Eulerian points of view can be observed with the *stochastic differential equation*

$$\begin{cases} dX_t(x) = b(t, X_t(x))dt + \sqrt{2\nu}dW_t \\ X_0(x) = x \end{cases} \tag{0.0.7}$$

and the *advection-diffusion* equation

$$\begin{cases} \partial_t u + \mathrm{div}_x(bu) = \nu\Delta_x u & \text{in } (0,T)\times\mathbb{T}^d \\ u(0,\cdot) = u_0 & \text{if } t = 0 \end{cases}, \tag{0.0.8}$$

since the solution u of (0.0.8) can be represented in terms of the solution X_t of (0.0.7) via the *Feynman-Kac* formula

$$u(t,\cdot) = \mathbb{E}[X(t,\cdot)_\# u_0],$$

where $X(t,\cdot)_\# u_0$ denotes the push-forward of the initial datum. The basic idea is the same as the method of characteristics: the initial datum is transported via the stochastic flow X_t. However, since the solution of (0.0.7) is not deterministic, an averaging procedure is required, which explains the expected value in the Feynman-Kac formula. The theory of well-posedness in low regularity settings for (0.0.7) and (0.0.8) is vast and we will not treat it in detail in this thesis. We only mention that the Brownian motion has a *very* strong regularising effect: heuristically, the presence of a Brownian perturbation prevents the formation of singular measures, which is exactly what we avoid by considering regular Lagrangian flows in (0.0.4). Well-posedness results for (0.0.7) (with $\nu > 0$) were already known before [70] and since the Brownian motion prevents the formation of singularities, no regularity on the vector field is actually needed. In particular, Veretennikov [156] proved existence and pathwise uniqueness for solutions of (0.0.7) only requiring the vector field b to be bounded. We point out that the connection between (0.0.7) and (0.0.8) still holds when the vector field b itself is random. In other words, a non-deterministic velocity field in (0.0.7) can have the effect, at the Eulerian level, of enhancing the diffusivity ν. Indeed, averaging over possible directions of advection produces an effect analogous to a random walk, which is well known to induce diffusion.

We now briefly describe the contents of the various chapters, highlighting the original contributions of the author. We refer the reader to the introduction of each chapter for more details.

In *Chapter 1* we review classical well-posedness theory for (0.0.1), (0.0.2) and (0.0.3),(0.0.4), showing how they are related and how usual assumptions often considered during introductory analysis courses can be relaxed.

In *Chapter 2* we present the results of DiPerna and Lions [70], establishing well-posedness at the PDE level. We also discuss the general theory for ODEs with non-regular vector fields by introducing the general framework developed by Ambrosio [7]. We conclude the chapter with a discussion on non-uniqueness phenomena and their construction, ranging from classical examples by Aizenman [3] and Depauw [66] to convex integration approaches [126, 127, 125], up to very recent results [37].

Chapter 3 is based on a paper by the author [55], where uniqueness of solutions of (0.0.1) in a low-regularity setting is proved without using the renormalization technique. Instead, a more geometric argument is employed, relying on currents from geometric measure theory. In particular, the key observation is to notice that (0.0.1) can be seen as a condition on the boundary of a 1-current of a specific form.

In *Chapter 4* we prove stability estimates and other results from [58]: the techniques used in the proofs will be the blueprint for the results in the subsequent chapters. We also present a novel quantitative stability estimate in a strong topology for the solutions of (0.0.1).

In *Chapter 5* we present the results obtained by the author in [56]. We devise a numerical scheme, based on the explicit Euler method, to approximate both the regular Lagrangian flow of a vector field and the solution of (0.0.1) in a low-regularity setting. In Section 5.5 we also present the results obtained by the author in collaboration with Gennaro Ciampa, Gianluca Crippa, Raffaele D'Ambrosio and Stefano Spirito in [46]. In particular, in [46] we prove approximation results for the regular Lagrangian flow (which we also use to prove approximation results at the PDE level) using a θ-method, and we provide numerical simulations which show the logarithmic convergence rate we predict.

In *Chapter 6*, which also presents mainly original results, we show how the techniques employed in Chapter 5 can be applied to the context of SDEs and the advection-diffusion equation, approximating the *stochastic flow* X_t defined in (0.0.7) with a Euler-Maruyama scheme and using the Feynman-Kac formula to get related approximation results for (0.0.8).

In *Chapter 7*, the focus is on mathematical modeling. Specifically, the objective of this chapter is to study of how advection induced by random forces, produced by microorganisms immersed in a fluid, gives rise to enhanced diffusion. The original contribution of the author, in collaboration with prof. Giovanni Alberti and prof. Antonio DeSimone, is presented in section 7.4, where we devise a mathematical model to predict the enhanced diffusion in a fluid produced by the action of an ensemble of Stokeslets above a no-slip floor. This study is motivated by the real-world microorganism *Stylonychia*, which feed in a similar fashion [155]. The results in this chapter are still a work in progress, which will be finalised in [6].

Notations

The main notations used in the thesis are listed below.

We denote by $\mathbb{R}^d$ the d-dimensional Euclidean space and by $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ the d-dimensional torus, each one considered as a metric space with their respective standard distances. We denote by $\mathbb{R}^+$ and $\mathbb{R}^+_0$ the sets of positive and non-negative real numbers, respectively. We also denote by I an interval, often we will have $I = [0, T)$ for some $T > 0$. The d-dimensional Lebesgue measure is denoted by $\mathscr{L}^d$, and we denote by $\mathcal{H}^d$ the d-dimensional Hausdorff measure. For other generic measures we will typically use the letters μ or ν. Given a measure space $(X, \mathcal{F}, \mu)$, with $\mathcal{F}$ a σ-algebra, we will denote by $\mathcal{M}(X)$ and $\mathcal{M}_+(X)$ the sets of measures and of positive measures on the measurable space $(X, \mathcal{F})$, respectively. If X is a topological space we will implicitly assume that $\mathcal{F}$ is the Borel σ-algebra, if not otherwise specified.

Given $\Omega \subset \mathbb{R}^d$ (or $\Omega \subset \mathbb{T}^d$), we denote by $L^p(\Omega)$ and $W^{1,p}(\Omega)$, with $1 \leq p < \infty$, the sets

$$L^p(\Omega) = \left\{ f : \Omega \to \mathbb{R} \text{ measurable and such that } \int_\Omega |f|^p dx < +\infty \right\}$$

and

$$W^{1,p}(\Omega) = \left\{ f \in L^p(\Omega) \text{ such that } \int_\Omega \|\nabla_x f\|^p dx < +\infty \right\}.$$

If $p = +\infty$, $L^\infty(\Omega)$ and $W^{1,\infty}(\Omega)$ denote the spaces of measurable functions such that $|f|$ is $\mathscr{L}^d$-almost everywhere bounded and such that $|f| + \|\nabla_x f\|$ is $\mathscr{L}^d$-almost everywhere bounded, respectively. We may consider generalisation of such spaces with vector-valued functions or with respect to different measures. We denote by b a vector field and by either X or Φ its flow. In Chapter 3 and Chapter 5 (and in the statement of some results) we shall use the letter Φ to be consistent with the reference material [55, 56], otherwise we use X. We use $\log(\cdot)$ to denote the natural logarithm.

We use personalised tags for some equations. For the reader's convenience, we list here these tags together with where they are defined in the text.

- (CE), (TE), (DP-L) and (ρ-TE) are defined in Chapter 2, Section 1;
- (μ-CE) and (ODE) are defined in Chapter 2, Section 2;
- (Hyp) and (C-CE) are defined in Chapter 3, Section 2;
- (PDE) and (MA) are defined at the beginning of Chapter 5 and in Chapter 5, Section 1, respectively;
- (ADE), (B-SDE), (SDE) and (F-K) are defined in Chapter 6, Section 1.

Chapter 1

The classical setting

1.1 The ODE theory: existence, uniqueness and stability

1.1.1 Classical results on ODEs

In this section we will prove some of the main results of the classical theory on ODEs and we will show how this theory can be applied to prove related results at the PDE level on the continuity and the transport equations. Even though such results are very standard and well known, we will give detailed proofs (or references where the interested reader can find detailed proofs) both for the reader's convenience and because many of the basic ideas used in the classical setting carry over, with due modifications, to the non-classical case we are going to discuss in the next chapters.

Let $b : I_{t_0} \times \mathbb{R}^d \to \mathbb{R}^d$ be a vector field, with I_{t_0} an open neighbourhood of $t_0 \in \mathbb{R}$. We consider the following Cauchy problem:

$$\begin{cases} \frac{d}{dt}\gamma(t) = b(t, \gamma(t)) \text{ if } t \in I_{t_0} \\ \gamma(t_0) = x_0 \end{cases} . \tag{1.1.1}$$

We will discuss existence, uniqueness and stability of classical solutions of (1.1.1). The following well-known example shows that uniqueness of solutions is not always guaranteed:

Example 1.1.1. ***The square root example***
Consider the following Cauchy problem on $\mathbb{R}$*:*

$$\begin{cases} \frac{d}{dt}\gamma(t) = 2\sqrt{|\gamma(t)|} \\ \gamma(0) = 0 \end{cases} . \tag{1.1.2}$$

It can be easily verified that, for every $c \geq 0$*, the following function is a solution for* (1.1.2)*:*

$$\gamma_c(t) = \begin{cases} 0 & \textit{if } t \leq c \\ (t-c)^2 & \textit{if } t > c. \end{cases}$$

We now start with a list of classical theorems dealing with existence and uniqueness of solutions of (1.1.1). We will give two versions of each theorem, showing that the continuity in time and the uniform Lipschitz condition usually assumed in such results are not really necessary. The reason we do this is to immediately show how the regularity in time is not the issue in these kinds of problems, especially when considering the uniqueness of solutions [1]. The first result on existence and uniqueness of solutions of (1.1.1) is one of the most famous theorems in classical analysis:

[1]Once we drop the continuity assumption in time and only assume summability, issues about the well-posedness of (1.1.1) in a pointwise sense arise. One has to consider a representative of the vector field which is defined pointwise.

Theorem 1.1.1. ***Cauchy-Lipschitz-Picard-Lindelöf existence and uniqueness theorem, [57, Theorem 1.2.2]***
Let b be a continuous vector field defined on an open set of $\mathbb{R} \times \mathbb{R}^d$ containing the rectangle

$$D = \{(t, x) \in \mathbb{R} \times \mathbb{R}^d \text{ such that } |t - t_0| \leq \alpha \text{ and } |x - x_0| \leq \beta\}.$$

Assume that b is uniformly Lipschitz continuous in the space variable with respect to time in D, i.e.

$$|b(t, x) - b(t, y)| \leq K|x - y| \quad \text{for every } (t, x), (t, y) \in D, \text{ for some } K > 0. \tag{1.1.3}$$

Moreover, assume that $|b(t, x)| \leq M$ in D for some $M > 0$. Then there exists a unique solution $\gamma \in C^1([t_0 - r, t_0 + r]; \mathbb{R}^d)$ to (1.1.1) *with initial condition $\gamma(t_0) = x_0$, where*

$$r < \min\left\{\alpha, \frac{\beta}{M}, \frac{1}{K}\right\}.$$

Proof. See the proof of Theorem 1.1.2. □

The next alternative version of the theorem, which does not require time continuity or a uniform Lipschitz constant, is proved in essentially the same way as the classical one [2].

Theorem 1.1.2. ***Cauchy-Lipschitz-Picard-Lindelöf existence and uniqueness theorem, revisited***
Let $I = [0, T)$ with $T > 0$ be a time interval and let b be a vector field such that $|b(t, x)| \leq M$ for some $M > 0$ for every $(t, x) \in I \times \mathbb{R}^d$. Assume that

$$|b(t, x) - b(t, y)| \leq K(t)|x - y| \text{ for some } K \in L^1(I) \text{ for every } (t, x), (t, y) \in I \times \mathbb{R}^d$$

and that T is sufficiently small such that

$$\int_0^T K(t)dt < 1.$$

Then, for every initial datum $x_0 \in \mathbb{R}^d$, there exists a unique absolutely continuous solution $\gamma : I \to \mathbb{R}^d$ solving (1.1.1).

Remark 1.1.1.
Notice that in this theorem we do not assume continuity in time of b. Moreover it is easy to obtain the result for every $T > 0$ by considering $\gamma(T - \varepsilon)$ a new initial point and applying the theorem iteratively.

Proof. Consider $x_0 \in \mathbb{R}^d$ fixed. We will show that there exists a unique continuous curve $\gamma : I \to \mathbb{R}^d$ such that

$$\gamma(t) = x_0 + \int_0^t b(s, \gamma(s))ds \text{ for every } t \in I,$$

which is the integral formulation of (1.1.1). In order to do that, we consider the metric space

$$X = \{\gamma \in C(I; \mathbb{R}^d) \text{ such that } |\gamma(t) - x_0| \leq MT \ \forall t \in I \text{ and } \gamma(0) = x_0\}$$

endowed with the supremum norm. Since X is a closed subspace of $C(I; \mathbb{R}^d)$, it is a complete metric space. We consider the map $\mathcal{L} : C(I; \mathbb{R}^d) \to C(I; \mathbb{R}^d)$ defined as

$$\mathcal{L}[\gamma](t) := x_0 + \int_0^t b(s, \gamma(s))ds,$$

and we notice that $\mathcal{L}$ maps X onto itself: given $\gamma \in X$, it holds that $\mathcal{L}[\gamma](0) = x_0$ and

$$|\mathcal{L}[\gamma](t) - x_0| = \left|\int_0^t b(s, \gamma(s))ds\right| \leq \int_0^t |b(s, \gamma(s))|ds \leq MT \text{ for every } t \in I.$$

[2] We will be interested mainly in initial value problems, considering an initial time $t = 0$ and going forward in time. So, instead of a neighbourhood of $t_0 \in \mathbb{R}$, we will consider a time interval $I = [0, T)$ with $T > 0$.

We now show that $\mathcal{L}$ is a contraction. Given $\gamma_1, \gamma_2 \in X$ it holds that for every $t \in I$:

$$|\mathcal{L}[\gamma_1](t) - \mathcal{L}[\gamma_2](t)| = \left| \int_0^t b(s,\gamma_1(s)) - b(s,\gamma_2(s)) ds \right| \le \int_0^t |b(s,\gamma_1(s)) - b(s,\gamma_2(s))| ds$$
$$\le \int_0^t K(s)|\gamma_1(s) - \gamma_2(s)| ds \le \left(\int_0^T K(t) dt \right) \sup_{t\in I} |\gamma_1(t) - \gamma_2(t)|.$$

Taking the supremum over $t \in I$ on the left-hand side:

$$\sup_{t\in I} |\mathcal{L}[\gamma_1](t) - \mathcal{L}[\gamma_2](t)| \le \underbrace{\left(\int_0^T K(t) dt \right)}_{<1} \sup_{t\in I} |\gamma_1(t) - \gamma_2(t)| < \sup_{t\in I} |\gamma_1(t) - \gamma_2(t)|.$$

This proves that $\mathcal{L}$ is a contraction on X and by Banach's fixed point theorem there exists a unique fixed point $\gamma \in X$, which is the unique solution of (1.1.1). The absolute continuity of γ simply follows by the absolute continuity of the integral.

□

If one drops the Lipschitz continuity assumption on the vector field (or the weaker assumption of Theorem 1.1.2) one loses uniqueness, but existence can still hold.

Theorem 1.1.3. ***Peano existence theorem , [57, Theorem 1.2.4]***
Let b be a continuous vector field defined on an open set containing the rectangle

$$D = \left\{ (t,x) \in \mathbb{R} \times \mathbb{R}^d \text{ such that } |t - t_0| \le \alpha \text{ and } |x - x_0| \le \beta \right\}$$

and such that $|b(t,x)| \le M$ in D. Then there exists a solution to (1.1.1) *in $C^1([t_0 - r, t_0 + r]; \mathbb{R}^d)$ with $r < \min \left\{ \alpha, \frac{\beta}{M} \right\}$.*

Similarly to what we considered previously, there is a stronger version of the classical Peano theorem where continuity in time is not required.

Theorem 1.1.4. ***Carathéodory existence theorem , [47, Theorem 1.1]***
Let b be a vector field defined on an open set containing the rectangle

$$D = \left\{ (t,x) \in \mathbb{R} \times \mathbb{R}^d \text{ such that } t_0 \le t \le t_0 + \alpha \le \alpha \text{ and } |x - x_0| \le \beta \right\}.$$

Assume that b satisfies, on D, the following Carathéodory conditions*:*

1. *$x \mapsto b(t,x)$ is continuous for almost every $t \in [t_0, t_0 + \alpha]$;*
2. *$t \mapsto b(t,x)$ is measurable in t for every x such that $|x - x_0| \le \beta$;*
3. *$|b(t,x)| \le m(t)$ for every $(t,x) \in D$, with $m(t) \in L^1([t_0, t_0 + \alpha])$.*

Then (1.1.1) *admits an absolutely continuous solution γ in the time interval $[t_0, t_0 + d]$, where $d \le \alpha$ is such that*

$$\int_{t_0}^{t_0+d} m(t) dt \le \beta.$$

Remark 1.1.2. ***Continuity in time: if it is not really necessary, why is it there?***
We think it is appropriate to add a brief remark on why the continuity in time is often assumed in results concerning existence and uniqueness of ODEs. There are two reasons in our opinion: one pedagogical and the other one historical. The Cauchy-Lipschitz-Picard-Lindelöf theorem and the Peano theorem are usually taught to students who have not been introduced to the Lebesgue integral yet, and historically such theorems need continuity in time because they originally used the Riemann integral rather than Lebesgue's. Indeed they were first proved in [118] and [135] respectively, and at that time the Lebesgue integral simply did not exist: it was invented later [112].

We now prove a version (one of the many) of the Grönwall's lemma [84], which will be very useful in proving some of the subsequent results.

Lemma 1.1.1. ***(A version of) Grönwall's lemma***
Let $I = [0, T)$ for some $T > 0$ and consider $f, g, C \in L^1(I; \mathbb{R}_0^+)$. If for almost every $t \in I$

$$f(t) \leq C(t) + \int_0^t f(s)g(s)ds, \tag{1.1.4}$$

then it holds that

$$f(t) \leq C(t) + \int_0^t \exp\left(\int_s^t g(z)dz\right) g(s)C(s)ds \quad \text{for almost every } t \in I. \tag{1.1.5}$$

If moreover C is non-decreasing, then

$$f(t) \leq C(t)\exp\left(\int_0^t g(s)ds\right) \quad \text{for almost every } t \in I. \tag{1.1.6}$$

Proof. Consider the function

$$h(t) := \exp\left(-\int_0^t g(s)ds\right)\int_0^t f(s)g(s)ds.$$

Differentiating h we get that for almost every $s \in [0, t]$:

$$\begin{aligned}
h'(s) &= -g(s)\exp\left(-\int_0^s g(z)dz\right)\left(\int_0^s f(z)g(z)dz\right) + \exp\left(-\int_0^s g(z)dz\right) f(s)g(s) \\
&= \exp\left(-\int_0^s g(z)dz\right) g(s)\left(f(s) - \int_0^s f(z)g(z)dz\right) \leq \exp\left(-\int_0^s g(z)dz\right) g(s)C(s).
\end{aligned}$$

Integrating the above inequality in $[0, t]$ and since $h(0) = 0$:

$$h(t) \leq \int_0^t \exp\left(-\int_0^s g(z)dz\right) g(s)C(s)ds.$$

By the definition of h:

$$\begin{aligned}
\int_0^t f(s)g(s)ds = h(t)\exp\left(\int_0^t g(s)ds\right) &\leq \exp\left(\int_0^t g(s)ds\right)\int_0^t \exp\left(-\int_0^s g(z)dz\right) g(s)C(s)ds \\
&= \int_0^t \exp\left(\int_0^t g(s)ds\right)\exp\left(-\int_0^s g(z)dz\right) g(s)C(s)ds \\
&= \int_0^t \exp\left(\int_s^t g(z)dz\right) g(s)C(s)ds.
\end{aligned} \tag{1.1.7}$$

Substituting $\int_0^t f(s)g(s)ds$ in (1.1.4) with the last term in (1.1.7) we get (1.1.5). If C is non-decreasing, then

$$\begin{aligned}
f(t) &\leq C(t) + \int_0^t \exp\left(\int_s^t g(z)dz\right) g(s)C(s)ds \leq C(t) + C(t)\int_0^t g(s)\exp\left(\int_s^t g(z)dz\right) ds \\
&= C(t) + C(t)\int_0^t \frac{d}{ds}\left[-\exp\left(\int_s^t g(z)dz\right)\right] ds = C(t) + C(t)\left[-\exp\left(\int_s^t g(z)dz\right)\Bigg|_{s=0}^{s=t}\right] \\
&= C(t) + C(t)\left[-1 + \exp\left(\int_0^t g(z)dz\right)\right] = C(t)\exp\left(\int_0^t g(z)dz\right),
\end{aligned}$$

and we have proved (1.1.6).

□

In the classical setting there are at least two conditions which are more general than the Lipschitz one required in Theorem 1.1.1 and that ensure uniqueness of the solution of (1.1.1). The first one is the *one-sided Lipschitz condition*.

Proposition 1.1.1. ***One-sided Lipschitz condition***
Uniqueness forward in time for (1.1.1) *(i.e. two solutions γ_1, γ_2 coincide for times $t \geq t_0$) holds if the Lipschitz condition* (1.1.3) *in Theorem 1.1.1 is replaced by the following* one-sided Lipschitz condition*:*

$$\langle b(t,x) - b(t,y), x - y\rangle \leq K|x-y|^2 \text{ for every } (t,x),(t,y) \in D,$$

with $\langle \cdot, \cdot \rangle$ denoting the standard scalar product in $\mathbb{R}^d$.

Proof. Assume there are two solutions γ_1, γ_2 for (1.1.1). Define $r(t) := |\gamma_1(t) - \gamma_2(t)|^2$. It holds that $r(t_0) = 0$ and

$$\begin{aligned}\frac{d}{dt} r(t) &= 2\langle \gamma_1(t) - \gamma_2(t), b(t,\gamma_1(t)) - b(t,\gamma_2(t))\rangle \\ &\leq 2K|\gamma_1(t) - \gamma_2(t)|^2 = 2Kr(t).\end{aligned}$$

Integrating the above inequality we have $r(t) \leq 2K \int_{t_0}^t r(s)ds$, and by Grönwall's lemma we conclude that $r(t) \equiv 0$.

□

Remark 1.1.3. ***Non-uniform one-sided Lipschitz constant***
As we extended the result of Theorem 1.1.1 to Theorem 1.1.2 assuming a time-dependent Lipschitz constant $K(t)$, the same generalisation applies here provided that $K \in L^1(I)$.

Remark 1.1.4. ***Backward uniqueness***
It is easy to see how assuming a lower bound like

$$\langle b(t,x) - b(t,y), x - y\rangle \geq -K|x-y|^2 \tag{1.1.8}$$

we can prove uniqueness also backward in time. Indeed, we can reduce the (backward) Cauchy problem

$$\begin{cases} \frac{d}{dt}\bar{\gamma}(t) = b(t,\bar{\gamma}(t)) \text{ if } t \in (-T,0) \\ \bar{\gamma}(0) = x_0 \end{cases}$$

to the (forward) Cauchy problem

$$\begin{cases} \frac{d}{dt}\gamma(t) = -b(-t,\gamma(t)) \text{ if } t \in (0,T) \\ \gamma(0) = x_0 \end{cases}$$

just by noticing that given a solution γ of the latter , $\bar{\gamma}(t) := \gamma(-t)$ solves the former, and uniqueness holds if $-b$ has the one-sided Lipschitz property, which is precisely condition (1.1.8).

Another sufficient condition that ensures uniqueness of the solution of (1.1.1) is the so-called *Osgood condition* [133].

Proposition 1.1.2. ***Osgood condition, [57, Proposition 1.2.7]***
Uniqueness for (1.1.1) *holds if the Lipschitz condition* (1.1.3) *in Theorem 1.1.1 is replaced by the following Osgood condition:*

$$|b(t,x) - b(t,y)| \leq \omega(|x-y|) \text{ for every } (t,x),(t,y) \in D,$$

where $\omega : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is a continuous non-decreasing function satisfying $\omega(0) = 0$, $\omega(z) > 0$ for every $z > 0$ and

$$\int_0^1 \frac{1}{\omega(z)} dz = +\infty.$$

Remark 1.1.5. ***Non-existence***
While the Osgood condition ensures the continuity in the spatial variables of b, and thus up to other suitable assumptions the existence of a C^1 solution by Theorem 1.1.4, this is not true for the one-sided Lipschitz condition since every monotone decreasing function (possibly discontinuous) satisfies it. For instance, consider

$$g(t) = \begin{cases} \pi & \text{if } t < 0 \\ e & \text{if } t \geq 0 \end{cases}$$

and the Cauchy problem

$$\begin{cases} y'(t) = g(y(t)) \\ y(0) = 0 \end{cases} .$$

A C^1 solution $y(t)$ does not exist since it can be easily verified that y is not differentiable in 0. However, if we consider for instance Lipschitz solutions instead of C^1 ones, it is easy to see that

$$y(t) = \begin{cases} \pi t & \text{if } t \leq 0 \\ et & \text{if } t > 0 \end{cases}$$

is a valid solution. Namely, for almost every t it holds that $y'(t) = g(y(t))$. This is not always the case: there are examples with g merely bounded where no Lipschitz solutions of (1.1.1) exist. We report here a proof proposed in [93], based on arguments from [87]. Let $A \subseteq \mathbb{R}$ be a Borel set of positive (but not full) measure in each interval. The existence of such a set is proved in [141]. Let $F(x) = \mathbb{1}_A(x) - \mathbb{1}_{A^c}(x)$ and assume that $\exists \varphi : (a,b) \to \mathbb{R}$ a Lipschitz integral curve for F. The set $R = \varphi((a,b))$ is an interval (by the intermediate value property) with positive Lebesgue measure (because $\varphi'(t) \neq 0$ for almost every $t \in (a,b)$). Consider the set

$$\tilde{R} = \{x \in R \text{ such that } \forall t \in \varphi^{-1}(\{x\})\ \exists \varphi'(t) = F(\varphi(t))\}.$$

The set $R \setminus \tilde{R}$ has measure zero, being image of the zero-measure set

$$N = \{t \in (a,b) \text{ such that } \nexists \varphi'(t) \text{ or } \varphi'(t) \neq F(\varphi(t))\}$$

under the Lipschitz map φ. Notice that for any $x \in \tilde{R}$ there exists just one $t \in (a,b)$ such that $\varphi(t) = x$. Indeed, otherwise one could find distinct points $t_1, t_2 \in (a,b)$ such that $\varphi(t_1) = \varphi(t_2) = x \in \tilde{R}$ and $\varphi(t) - x$ has constant sign on (t_1, t_2). Then $\varphi'(t_1)$ and $\varphi'(t_2)$ have different signs, a contradiction with the definition of $\tilde{R}$. Therefore φ is injective on $\varphi^{-1}(\tilde{R})$ and it is easy to show that φ is injective on the whole (a,b), using the fact that $\varphi'(t) \neq 0$ for almost every $t \in (a,b)$. By continuity and injectivity the function φ is strictly monotone, hence one of the sets $R \cap A$ or $R \setminus A$ has measure zero (because φ cannot map a negligible set to a set with positive measure). This contradicts the definition of A.

Remark 1.1.6. ***The case*** $b \notin L^\infty$
Although we will assume vector fields to be bounded in order to avoid technical complications we remark that the typical assumption needed when working with unbounded vector fields is that $|b|$ is dominated by a linear function. Without delving too much in technicalities, the issue is that if b grows superlinearly a blow-up in finite time might occur. An example, assuming $b(x) = x^2$, is the solution of

$$\begin{cases} \frac{d}{dt} y(t) = y(t)^2 \\ y(0) = 1 \end{cases} ,$$

which is $y(t) = -\frac{1}{t-1}$. A similar requirement is needed also in low-regularity settings [70, 7].

1.1.2 Classical flow of a vector field

We now want to compare solutions γ_1, γ_2 of (1.1.1) with different initial conditions $\gamma_1(0) = x_1$, $\gamma_2(0) = x_2$ and different driving vector fields b_1, b_2 (we will always assume $t_0 = 0$ and we will consider solutions in $I = [0, T)$ for some $T > 0$). For this reason, it is convenient to introduce the notion of *flow* of a vector field.

Definition 1.1.1. ***(Classical) flow of a vector field***
Let $b : I \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous and bounded vector field. The (classical) flow of b is a map

$$X : I \times \mathbb{R}^d \to \mathbb{R}^d$$

such that it solves

$$\begin{cases} \partial_t X(t,x) = b(t, X(t,x)) \\ X(0,x) = x \end{cases} \quad \text{for every } (t,x) \in I \times \mathbb{R}^d,$$

i.e.

$$X(t,x) = x + \int_0^t b(s, X(s,x))ds \quad \text{for every } (t,x) \in I \times \mathbb{R}^d.$$

If b is Lispchitz continuous in the spatial variables (or if it satisfies the weaker assumptions of Theorem 1.1.2) we have existence and uniqueness of the flow X, since for every $(t,x) \in I \times \mathbb{R}^d$ we have that

$$X(t,x) = \gamma(t), \text{ with } \gamma \text{ the unique solution of } \begin{cases} \frac{d}{dt}\gamma(t) = b(t,\gamma(t)) \\ \gamma(0) = x \end{cases}.$$

We will now prove that under suitable assumptions the map X is Lipschitz continuous with respect to both the time t and the initial datum x.

Proposition 1.1.3. ***Stability properties of the classical flow***
Let X_1, X_2 be classical flows, as defined in Definition 1.1.1, with respect to bounded and uniformly (in time) Lipschitz (in space) vector fields b_1 and b_2. Let L_1 and L_2 be the Lipschitz constants of b_1 and b_2 respectively, and let $L := \min\{L_1, L_2\}$ [3]. It holds that, for every $(t,x),(t,y) \in I \times \mathbb{R}^d$:

- *If $b_1 = b_2 = b$, then $X_1 = X_2 = X$ and*

$$|X(t,x) - X(s,x)| \le \|b\|_{L^\infty}|t-s| \text{ and } |X(t,x) - X(t,y)| \le \min\{|x-y|e^{Lt}, |x-y| + 2\|b\|_\infty t\}.$$

- *If $x = y$, then*

$$|X_1(t,x) - X_2(t,x)| \le t(\|b_1\|_{L^\infty} + \|b_2\|_{L^\infty})e^{Lt}.$$

- *Assuming $s \le t$, it holds*

$$|X_1(t,x) - X_2(s,y)| \le \|b_1\|_{L^\infty}|t-s| + |x-y|e^{L_1 s} + s\|b_1 - b_2\|_{L^\infty}e^{Ls}$$

and

$$|X_1(t,x) - X_2(s,y)| \le \|b_1\|_{L^\infty}|t-s| + |x-y| + s(\|b_1\|_{L^\infty} + \|b_2\|_{L^\infty}).$$

Proof. First, let us assume that $b_1 = b_2 = b$. In this case $X_1 = X_2 = X$ and:

$$|X(t,x)-X(t,y)| \le |x-y| + \int_0^t |b(s,X(s,x))-b(s,X(s,y))|ds \le |x-y| + L\int_0^t |X(s,x)-X(s,x)|ds.$$

By Grönwall's lemma it follows that

[3] As it will be clear from the proof, the main tool is the Grönwall's lemma. Thus, there are easy generalisations possible, for instance analogous estimates hold requiring that the Lipschitz constant $L(t)$ varies in time and $L \in L^1(I)$. In particular it is not necessary for L_1 and L_2 to be bounded.

$$|X(t,x)-X(t,y)| \leq |x-y|e^{Lt}.$$

It is trivial to see that $|X(t,x)-X(s,x)| \leq \|b\|_{L^\infty}|t-s|$ and $|X(t,x)-X(t,y)| \leq |x-y|+2\|b\|_{L^\infty}t$. Similarly, assuming $x=y$ but $b_1 \neq b_2$ we get:

$$\begin{aligned}
&|X_1(t,x)-X_2(t,x)| \leq \int_0^t |b_1(s,X_1(s,x))-b_2(s,X_2(s,x))|ds \\
&\leq \int_0^t |b_1(s,X_1(s,x))-b_2(s,X_1(s,x))|+|b_2(s,X_1(s,x))-b_2(s,X_2(s,x))|ds \\
&\leq t\|b_1-b_2\|_{L^\infty}+L_2\int_0^t |X_1(s,x)-X_2(s,x)|ds.
\end{aligned}$$

Since the argument above is symmetric in b_1, b_2, we have by Grönwall's lemma:

$$|X_1(t,x)-X_2(t,x)| \leq t\|b_1-b_2\|_{L^\infty}e^{Lt}$$

with $L=\min\{L_1,L_2\}$. We finally combine the two estimates to compare X_1 and X_2 if $x \neq y$ and $b_1 \neq b_2$ at different times. Assuming that $s \leq t$ it is sufficient to consider a third flow X_3 with respect to the vector field b_1 and with initial datum $X_3(0,y)=y$. Then:

$$\begin{aligned}
|X_1(t,x)-X_2(s,y)| &\leq |X_1(t,x)-X_1(s,x)|+|X_1(s,x)-X_3(s,y)|+|X_3(s,y)-X_2(s,y)| \\
&\leq \|b_1\|_{L^\infty}|t-s|+|x-y|e^{L_1 s}+s\|b_1-b_2\|_{L^\infty}e^{Ls}.
\end{aligned}$$

Finally, we notice that since we assumed b_1 and b_2 to be bounded it holds another trivial estimate, namely:

$$|X_1(t,x)-X_2(s,y)| \leq \|b_1\|_{L^\infty}|t-s|+|x-y|+s(\|b_1\|_{L^\infty}+\|b_2\|_{L^\infty}).$$

□

If b is more regular in space, this regularity is carried over to the associated flow. Namely, the following holds:

Proposition 1.1.4. ***C^1 regularity of the classical flow, [92, Theorem 3.1]***
Let $b : I \times \mathbb{R}^d \mapsto \mathbb{R}^d$ be a bounded vector field such that $b(t,\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is C^1 uniformly with respect to time. Then, given the flow X associated to b, it holds that $X(t,\cdot)$ is also C^1 for every $t \in I$.

The above Proposition can be iterated, proving that if b has spatial C^k regularity, uniformly in time, then also the flow X has such a C^k regularity in space:

Corollary 1.1.1. ***Higher order regularity of the classical flow, [92, Theorem 4.1]***
Let $b : I \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded smooth vector field. Then for every $t_0 \in I$ there exists a unique classical flow X with $X(t_0,x)=x$ for every $x \in \mathbb{R}^d$ that is smooth in both time and space.

We now introduce the notion of *backward flow*, which will be relevant in the proof of the next proposition and in the next sections.

Definition 1.1.2. ***Backward flow***
Let $b : I \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous and bounded vector field. For every $t \in I$ and for every $x \in \mathbb{R}^d$, we define the backward flow $Y_t : [0,t] \times \mathbb{R}^d \to \mathbb{R}^d$, denoted by $Y_t(s,y)$, as the solution of

$$\begin{cases} \partial_s Y_t(s,y) = -b(t-s,Y_t(s,y)) & \text{for } s \in [0,t] \\ Y_t(0,y)=y. & \text{for } s=0 \end{cases}.$$

It is important to remark that the above definition is motivated by the fact that (as we will prove in the next proposition) for every $t \in I$ the map $X(t,\cdot)$ is a diffeomorphism and $Y_t(t,\cdot) = X^{-1}(t,\cdot)$. The idea is the following: we know that given a flow X, for each $x \in \mathbb{R}^d$ by definition $X(t,x) = \gamma(t)$, with γ the unique integral curve solving (1.1.1) with $\gamma(0) = x$. If we now consider a point $y = \gamma(t_0)$, what the backward flow $Y_t(s,y)$ does is to transport back the point $y = \gamma(t_0)$ by following the curve γ in the opposite direction for a time s. Namely, it holds that

$$Y_t(s, X(t,x)) = X(t-s,x) \text{ for every } s \le t,\ x \in \mathbb{R}^d.$$

Proposition 1.1.5. ***Some properties of the classical flow***
Let $b : I \times \mathbb{R}^d \to \mathbb{R}^d$ be a smooth and bounded vector field. Then, given the (unique) flow X of b starting at $t = 0$, it holds that:

1. *The flow $X(t,x)$ satisfies a semigroup property, namely $X(t+s,x) = X(t,X(s,x)) = X(s,X(t,x))$ for every $x \in \mathbb{R}^d$ and for every $s,t \in I$ such that $t+s \le T$.*

2. *For every $t \in I$, the map $X(t,\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth diffeomorphism.*

3. *The Jacobian $J(t,x) = \det(\nabla_x X(t,x))$ satisfies the following equality:*

$$J(t,x) = \exp\left(\int_0^t \operatorname{div}_x(b)(s,X(s,x))ds\right) \tag{1.1.9}$$

with $\operatorname{div}_x(b)$ being the divergence (with respect to the spatial variables) of b. In particular, for every $(t,x) \in I \times \mathbb{R}^d$ it holds:

$$\exp\left(-\int_0^t \|\operatorname{div}_x(b)^-\|_{L^\infty} ds\right) \le J(t,x) \le \exp\left(\int_0^t \|\operatorname{div}_x(b)^+\|_{L^\infty} ds\right),$$

where $\operatorname{div}_x(b)^-$ and $\operatorname{div}_x(b)^+$ denote the negative and the positive parts of $\operatorname{div}_x(b)$, respectively.

Proof. The first point is trivial, and it follows by the uniqueness, for every $x \in \mathbb{R}^d$, of the solution of (1.1.1). The second point is also proved by using the uniqueness of the flow. Fix $t \in I$, $x \in \mathbb{R}^d$ and consider the backward flow $Y_t : [0,t] \times \mathbb{R}^d \to \mathbb{R}^d$, which is defined as

$$Y_t(s,y) := \gamma_{t,y}(s),$$

with $\gamma_{t,y}$ the (unique) solution of

$$\begin{cases} \partial_s \gamma_{t,y}(s) = -b(t-s,\gamma_{t,y}(s)) \\ \gamma_{t,y}(0) = y \end{cases}. \tag{1.1.10}$$

Since for every $y \in \mathbb{R}^d$ the solution of (1.1.10) is unique, we can easily see that if $y = X(t,x)$ for some $x \in \mathbb{R}^d$, then $Y_t(s,y) = \gamma_{t,y}(s) = X(t-s,x)$, by checking that $X(t-s,x)$ solves (1.1.10). Then, for every $x \in \mathbb{R}^d$:

$$Y_t(t, X(t,x)) = X(0,x) = x.$$

In the same way we see that

$$X(s, Y_t(t,y)) = Y_t(t-s,y),$$

thus

$$X(t, Y_t(t,y)) = Y_t(0,y) = y.$$

Then $X(t,\cdot)$ is invertible with inverse $Y_t(t,\cdot)$ and it is a diffeomorphism by the smoothness of Y_t, which follows by the smoothness of $-b(t-\cdot,\cdot)$ and Corollary 1.1.1. As for the final point, by Jacobi's formula [122, Chapter 8] (we remark that $\nabla_x X(t,x)$ is invertible since $X(t,\cdot)$ is a diffeomorphism) it holds:

$$\frac{d}{dt}J(t,x) = J(t,x)\,\mathrm{tr}\left(\nabla_x X(t,x)^{-1}\frac{d}{dt}[\nabla_x X(t,x)]\right),$$

with $\mathrm{tr}(A)$ being the trace of a matrix A. The flow is smooth map, so we can exchange the order of the derivatives to get

$$\frac{d}{dt}[\nabla_x X(t,x)] = \nabla_x\left[\frac{d}{dt}X(t,x)\right] = \nabla_x[b(t,X(t,x)] = \nabla_x b(t,X(t,x))\nabla_x X(t,x).$$

Since the trace of a matrix is invariant under similarity, it holds that

$$\begin{aligned}&\mathrm{tr}\left(\nabla_x X(t,x)^{-1}\frac{d}{dt}[\nabla_x X(t,x)]\right) = \mathrm{tr}\left(\nabla_x X(t,x)^{-1}\nabla_x b(t,X(t,x))\nabla_x X(t,x)\right)\\ &= \mathrm{tr}\left(\nabla_x b(t,X(t,x))\right) = \mathrm{div}_x\, b(t,X(t,x)).\end{aligned}$$

The formula (1.1.9) is thus proved, while the upper and lower bounds on $J(t,x)$ follow easily by Grönwall's lemma. □

Remark 1.1.7. ***Lipschitz vector fields instead of smooth ones***
In the above Proposition, a Lipschitz vector field (in the spatial variable only) and thus a Lipschitz flow would have been enough to prove an analogous result, with the due modifications.

1.2 The PDE theory: continuity equation and transport equation

We now turn our attention to PDEs which are strictly related to the Cauchy problem (1.1.1), as we will see shortly. We will denote with I an interval $I = [0,T)$ with $T > 0$. Given a bounded and smooth vector field $b : I \times \mathbb{R}^d \to \mathbb{R}^d$ and bounded smooth initial data u_0, ρ_0 [4], we say that $u : I \times \mathbb{R}^d \to \mathbb{R}$ is a solution of the *continuity equation* if:

$$\begin{cases}\partial_t u + \mathrm{div}_x(bu) = 0 & \text{in } \mathbb{R}^d \text{ with } t > 0\\ u(0,\cdot) = u_0 & \text{for t=0}\end{cases}, \tag{1.2.1}$$

while we say that $\rho : I \times \mathbb{R}^d \to \mathbb{R}^d$ is a solution of the *transport equation* if

$$\begin{cases}\partial_t \rho + b\cdot\nabla_x \rho = 0 & \text{in } \mathbb{R}^d \text{ with } t > 0\\ \rho(0,\cdot) = \rho_0 & \text{for t=0}\end{cases}. \tag{1.2.2}$$

Before proving existence and uniqueness of solutions, we provide a heuristic description of the scenarios represented by these two equations [5]. This will give both an intuition on the link between the equations and the Cauchy problem (1.1.1) and it will make the formulas for the solutions of (1.2.1) and (1.2.2) easier to interpret. We start with the continuity equation (1.2.1). This equation describes the change of *mass* in a certain region $\Omega \subseteq \mathbb{R}^d$, where we consider the mass in Ω at a time t to be $\int_\Omega u(t,x)dx$. If u satisfies (1.2.1), then by the divergence theorem:

$$\frac{d}{dt}\int_\Omega u(t,x)dx = \int_\Omega \frac{d}{dt}u(t,x)dx = -\int_\Omega \mathrm{div}_x(b(t,x)u(t,x))dx = \int_{\partial\Omega} u(t,x)\langle b(t,x),\nu(x)\rangle d\mathcal{H}^{d-1}(x),$$

with ν being the inward normal vector to $\partial\Omega$. This is telling us that the rate of change of the mass in Ω depends on the flux of b across $\partial\Omega$. It is really convenient to think of u as non-negative and describing an actual physical quantity (for instance, the density of a gas inside some region Ω). If the vector field b points inside Ω, i.e. $\langle b(t,x),\nu(x)\rangle > 0$, then it is "pushing" the mass $u(t,x)$ at the boundary of Ω inside of Ω, increasing the total mass inside Ω. If instead $\langle b(t,x),\nu(x)\rangle < 0$, it is pushing mass from inside Ω to the

[4] As we showed in the previous section, even in the classical setting smoothness and boundedness of the vector field are not really required. The same holds here, but in order to simplify the exposition we will assume u_0, ρ_0 and b to be smooth and bounded.

[5] We already briefly mentioned this difference in the Introduction, but we think it is really worth repeating it here, giving more details.

outside, decreasing the total mass, while if $\langle b(t,x), \nu(x)\rangle = 0$ the vector field is tangential to $\partial\Omega$ and the mass is conserved. It is worth noting that, since $u(t,x)$ multiplies the flux across $\partial\Omega$, if we assume $u_0 \geq 0$ we cannot end up with $u < 0$ in Ω. Moreover, if we consider a set Ω with null boundary the total mass is always conserved. This physical interpretation is telling us that the continuity equation is dealing with *densities* (or *measures* more generally), telling us how they evolve in time inside a region of space Ω. The transport equation (1.2.2), instead, has a different physical meaning. It describes the evolution of scalar quantities which are transported by the vector field b, but whose values remain unchanged. For instance, consider a handful of sand. We assume the initial datum to be

$$\rho_0(x) = \begin{cases} 1 & \text{if } x \text{ is a grain of sand} \\ 0 & \text{otherwise} \end{cases}.$$

If we throw the sand into a river, assuming that the grains of sand follow exactly the flow of the water, the evolution of ρ_0 will follow the transport equation. In this sense, the solution $\rho(t,x)$ is not a density whose value can increase or decrease, but rather $\rho(t,\cdot)$ is simply a "rearrangement" of the initial datum ρ_0. In the example above, the values of $\rho(t,x)$ will always be either 1 if there is a grain of sand at time t in position x or 0 if there is not. It is then easy to be convinced that if we are dealing with initial data which are absolutely continuous functions with respect to the Lebesgue measure and we assume $\Omega = \mathbb{R}^d$, the transport equation preserves both the L^1 norm (i.e. the total mass) and the L^∞ norm, while the continuity equation preserves the L^1 norm [6]. We now prove existence and uniqueness of solutions of (1.2.1) and (1.2.2). As we will see, having previously developed a Lagrangian theory on ODEs and flows will be crucial.

Proposition 1.2.1. ***Existence and uniqueness of solutions of*** (1.2.1) ***and*** (1.2.2)
Let $b : I \times \mathbb{R}^d \to \mathbb{R}^d$ be a smooth and bounded vector field and let u_0, ρ_0 be bounded smooth initial data for (1.2.1) *and* (1.2.2)*, respectively. Then, both* (1.2.1) *and* (1.2.2) *admit a unique solution in the class $C^\infty(I \times \mathbb{R}^d)$.*

Proof. We start by proving existence and uniqueness of solutions for (1.2.2). We consider Y_t the backward flow associated to b introduced in Definition 1.1.2. We recall that $Y_t(t,\cdot) = X^{-1}(t,\cdot)$, so by using the identity $X(t, Y_t(t,x)) = x$ and the chain rule we easily find that

$$\frac{d}{dt}[X(t,Y_t(t,x))] = 0 \text{ and } \nabla_x[X(t,Y_t(t,x))] = I.$$

This implies:

$$\frac{d}{dt}Y_t(t,x) = -[\nabla_x X(t,Y_t(t,x))]^{-1} b(t,x) \text{ and } \nabla_x Y_t(t,x) = [\nabla_x X(t,Y_t(t,x))]^{-1},$$

so

$$\frac{d}{dt}Y_t(t,x) = -\nabla_x Y_t(t,x)\cdot b(t,x).$$

At this point the existence follows by simply checking that the function

$$\rho(t,x) := \rho_0(Y_t(t,x)) \tag{1.2.3}$$

solves (1.2.2). To prove uniqueness, by the linearity of the equation we just need to prove that the solution $\rho \equiv 0$ is the unique solution if the initial datum is $\rho_0 = 0$. We argue by contradiction: if this were not the case, there would be a point $(t_0, x_0) \in I \times \mathbb{R}^d$ such that $\rho(t_0, x_0) \neq 0$. Consider now the backward curve $\gamma := Y_{t_0}(\cdot, x_0)$. We claim that the value of $\phi(s) := \rho(t_0 - s, \gamma(s))$ is constant. Indeed:

$$\frac{d}{ds}\phi(s) = -\partial_s\rho(t_0 - s,\gamma(s)) - \nabla_x\rho(t_0 - s,\gamma(s))\cdot b(t_0 - s,\gamma(s)) = 0.$$

Thus $\rho(t_0,x_0) = \phi(0) = \phi(t_0) = \rho(0,\gamma(t_0)) = 0$, a contradiction. As for the continuity equation, the existence of a solution is proved by checking that

[6] Beware that the heuristic discussion we just had is, indeed, only heuristic. We will see in later chapters how, with non-regular vector fields b, we will be able to find counterexamples where we have non zero solutions of (1.2.1) , (1.2.2) with a null initial datum.

$$u(t,x) := \frac{u_0(Y_t(t,x))}{\det(\nabla_x X(t,Y_t(t,x)))} \tag{1.2.4}$$

solves (1.2.1)[7]. By the linearity of the equation, the uniqueness is proved if we prove that the unique solution of (1.2.1) with $u_0 = 0$ as initial datum is $u = 0$. Assume that it is not, then we can assume without loss of generality that $\exists(t_0, x_0)$ such that $u(t_0, x_0) > 0$ and we can consider a ball $B_r(x_0)$ such that $u(t_0, x) > 0$ for every $x \in B_r(x_0)$. Consider the function

$$E(t) = \int_{B_{r(t)}(x_0)} u^2(t,x)dx \quad \text{with } r(t) = r + (t_0 - t)\|b\|_{L^\infty} \text{ and } t \in [0, t_0].$$

By Reynolds transport theorem [86]:

$$\begin{aligned}\frac{d}{dt}E(t) &= \int_{B_{r(t)}(x_0)} 2u(t,x)\frac{d}{dt}u(t,x)dx - \int_{\partial B_{r(t)}(x_0)} u^2(t,x)\|b\|_{L^\infty} d\mathcal{H}^{d-1}(x)\\ &= -\int_{B_{r(t)}(x_0)} 2u(t,x)\operatorname{div}_x(u(t,x)b(t,x))dx - \int_{\partial B_{r(t)}(x_0)} u^2(t,x)\|b\|_{L^\infty} d\mathcal{H}^{d-1}(x).\end{aligned}$$

We can check that

$$2u\operatorname{div}_x(bu) = \operatorname{div}_x(bu^2) + \operatorname{div}_x(b)u^2,$$

so

$$\begin{aligned}\frac{d}{dt}E(t) &= -\int_{B_{r(t)}(x_0)} \operatorname{div}_x(b(t,x))\, u^2(t,x) + \operatorname{div}_x(b(t,x)u^2(t,x))dx - \int_{\partial B_{r(t)}(x_0)} u(t,x)^2\|b\|_{L^\infty} d\mathcal{H}^{d-1}(x)\\ &= \int_{B_{r(t)}(x_0)} -\operatorname{div}_x(b(t,x))u^2(t,x)dx + \int_{\partial B_{r(t)}(x_0)} u^2(t,x)\underbrace{(\langle b(t,x)\cdot \nu(x)\rangle - \|b\|_{L^\infty})}_{\le 0} d\mathcal{H}^{d-1}(x)\\ &\le \|\operatorname{div}_x(b)^-\|_{L^\infty(K)}E(t),\end{aligned}$$

where K is a $(d+1)$-dimensional truncated cone in $\mathbb{R}_t\times\mathbb{R}_x^d$ with bases $B_{r+\|b\|_{L^\infty}t_0}(x_0)$ at $t = 0$ and $B_r(x_0)$ at $t = t_0$, and ν is the inward normal to $\partial B_{r(t)}(x_0)$. Since b is smooth and K is compact, $\|\operatorname{div}_x(b)^-\|_{L^\infty(K)} \le C$ for some $C > 0$, thus

$$\frac{d}{dt}E(t) \le CE(t).$$

Together with the fact that $E(0) = \int_{B_{r+\|b\|_{L^\infty}t_0}(x_0)} u_0^2(x)dx = 0$, by Grönwall's lemma this implies that

$$E(t_0) = \int_{B_r(x_0)} u^2(t,x)dx = 0,$$

a contradiction. □

Remark 1.2.1. ***Non-homogeneous equations and stability***
The existence and uniqueness of solutions of (1.2.1) *and* (1.2.2) *can also be proved when the right-hand side is not* 0*, but some function* f*. Using the backward flow and Duhamel's principle we have that*

$$u(t,x) = \frac{u_0(Y_t(t,x))}{\det[\nabla_x X(t,Y_t(t,x))]} + \int_0^t \frac{f(s,Y_t(t-s,x))}{\det[\nabla_x X(t,Y_t(t-s,x))]}ds$$

and

$$\rho(t,x) = \rho_0(Y_t(t,x))) + \int_0^t f(s,Y_t(t-s,x))ds$$

are solutions of (the inhomogeneous versions of) (1.2.1) *and* (1.2.2).

[7]See Proposition 2.2.1 for the interpretation of such a formula as the push-forward of a measure.

Chapter 2

The non-classical setting

2.1 The PDE theory: DiPerna-Lions renormalization technique

We start considering the well-posedness of the continuity and of the transport equations with non-regular vector fields. For the moment, we focus on proving existence, leaving the (considerably harder) issue of uniqueness for later. We will work with bounded vector fields defined on the d-dimensional torus $\mathbb{T}^d$. There are two reasons for which we make these assumptions:

1. We do not want to deal with the additional assumptions and technicalities needed when b is not bounded (see Remark 1.1.6).
2. Working in a compact manifold without boundary, we do not have to localise our arguments, avoiding additional technicalities in the proofs.

With the necessary additional assumptions if $b \notin L^\infty$ and with the due care if we work in $\mathbb{R}^d$ rather than $\mathbb{T}^d$, all the arguments we are proposing can be carried over. As in the previous chapter, we will denote with $I = [0, T)$ with $T > 0$ a time interval. At the cost of being redundant, we recall the continuity equation and the transport equation in $\mathbb{T}^d$, which are respectively

$$\begin{cases} \partial_t u + \mathrm{div}_x(bu) = 0 & \text{in } \mathbb{T}^d \text{ with } t > 0 \\ u(0, \cdot) = u_0 & \text{for t=0} \end{cases} \tag{CE}$$

and

$$\begin{cases} \partial_t \rho + b \cdot \nabla_x \rho = 0 & \text{in } \mathbb{T}^d \text{ with } t > 0 \\ \rho(0, \cdot) = \rho_0 & \text{for t=0} \end{cases} . \tag{TE}$$

From this point on, we will always consider *weak solutions* of (CE) and (TE). That is, we say that u solves (CE) if [1]

$$\int_0^T \int_{\mathbb{T}^d} u(\partial_t \varphi + b \cdot \nabla_x \varphi) dx dt + \int_{\mathbb{T}^d} \varphi(0, x) u_0(x) dx = 0 \quad \text{for every } \varphi \in C_c^\infty(I \times \mathbb{T}^d), \tag{2.1.1}$$

where $C_c^\infty(I \times \mathbb{T}^d)$ denotes the class of smooth functions with compact support in $I \times \mathbb{T}^d$. Similarly, ρ solves (TE) if

$$\int_0^T \int_{\mathbb{T}^d} \rho(\partial_t \varphi + \mathrm{div}_x(b)\varphi + b \cdot \nabla_x \varphi) dx dt + \int_{\mathbb{T}^d} \varphi(0, x) \rho_0(x) dx = 0 \quad \text{for every } \varphi \in C_c^\infty(I \times \mathbb{T}^d). \tag{2.1.2}$$

Remark 2.1.1.
We remark that from this point on we will always consider weak solutions, thus we will often simply write "solutions" instead of "weak solutions". When considering different kinds of solutions such as renormalized solutions or Lagrangian solutions, it will be specified.

[1] We omit explicitly writing the derivation of the formulas for the weak solutions, since it is standard.

We now choose the class where we will work, a choice that, as we shall see, will be crucial for establishing well-posedness. In the previous section the hypotheses on the vector field and on the initial datum were

$$\begin{cases} b \in [C^\infty(I \times \mathbb{R}^d) \cap L^\infty(I \times \mathbb{R}^d)]^d \\ u_0, \rho_0 \in C^\infty(\mathbb{R}^d) \end{cases}$$

and we proved existence and uniqueness of (strong) solutions in the class $C^\infty(I \times \mathbb{R}^d)$. In this section we will prove existence and uniqueness of solutions in the class

$$u, \rho \in L^\infty(0, T; L^q(\mathbb{T}^d)),$$

where $b \in [L^1(0, T; W^{1,p}(\mathbb{T}^d)) \cap L^\infty(I \times \mathbb{R}^d)]^d$ with an additional condition on $\mathrm{div}_x(b)$ we will specify shortly and $p, q \in [1, +\infty]$ such that $1/p + 1/q \leq 1$. First of all, let us show that we have existence of solutions in the considered class, even without a Sobolev regularity assumption on b:

Proposition 2.1.1. ***Existence of solutions***
Let $b \in [L^\infty((0, T) \times \mathbb{T}^d)]^d$ and $u_0, \rho_0 \in L^q(\mathbb{T}^d)$ with $q \in [1, +\infty]$. It holds that:

1. *If $\mathrm{div}_x(b)^- \in L^1(0, T; L^\infty(\mathbb{T}^d))$, there exists a solution of* (CE) *in $L^\infty(0, T; L^q(\mathbb{T}^d))$;*
2. *If $\mathrm{div}_x(b)^+ \in L^1(0, T; L^\infty(\mathbb{T}^d))$, there exists a solution of* (TE) *in $L^\infty(0, T; L^q(\mathbb{T}^d))$. If $q = +\infty$, we do not need this assumption on $\mathrm{div}_x(b)^+$. If $q = 1$, existence holds assuming also $\mathrm{div}_x(b)^- \in L^1(0, T; L^\infty(\mathbb{T}^d))$, i.e. $\mathrm{div}_x(b) \in L^1(0, T; L^\infty(\mathbb{T}^d))$.*

Proof. We mollify, in spatial variables only, the vector field b at scale $\delta > 0$ obtaining a smooth (in the spatial variables) vector field b^δ. For every $\delta > 0$, by Proposition 1.2.1 there exist solutions for (CE) and (TE) given by the explicit formulae (1.2.4) and (1.2.3). For the time being, we assume $1 < q < +\infty$. We start with the continuity equation. A solution, for every $\delta > 0$, is:

$$u^\delta(t, x) = \frac{u_0(Y_t^\delta(t, x))}{\det(\nabla_x X^\delta(t, Y_t^\delta(t, x)))},$$

with X^δ and Y^δ the forward and backward flow, respectively, relative to b^δ. We change variables setting $y = Y_t^\delta(t, x)$, so (since $Y_t^\delta(t, \cdot) = (X^\delta)^{-1}(t, \cdot)$) we have:

$$\int_{\mathbb{T}^d} |u^\delta(t, x)|^q dx = \int_{\mathbb{T}^d} \left| \frac{u_0(Y_t^\delta(t, x))}{\det(\nabla_x X^\delta(t, Y_t^\delta(t, x)))} \right|^q dx$$
$$= \int_{\mathbb{T}^d} \left| \frac{u_0(y)}{\det(\nabla_y X^\delta(t, y))} \right|^q \frac{dy}{|\det(\nabla_y Y_t^\delta(t, X^\delta(t, y)))|}.$$

Since $Y_t^\delta(t, X^\delta(t, y)) = y$, by the chain rule and Binet theorem

$$|\det(\nabla_y Y_t^\delta(t, X^\delta(t, y)))|^{-1} = |\det(\nabla_y X^\delta(t, y))|.$$

Substituting in the expression above:

$$\int_{\mathbb{T}^d} |u^\delta(t, x)|^q dx = \int_{\mathbb{T}^d} |u_0(y)|^q \frac{1}{|\det(\nabla_y X^\delta(t, y))|^{q-1}} dy. \tag{2.1.3}$$

By the assumption on the divergence and Proposition 1.1.5, we have that $1/|\det(\nabla_y X^\delta(t, y))|^{q-1}$ is uniformly bounded. So the sequence $\{u^\delta\}_{\delta>0}$ is uniformly bounded in $L^\infty(0, T; L^q(\mathbb{T}^d))$, and by extracting a subsequence we have weak-$*$ convergence to some $u \in L^\infty(0, T; L^q(\mathbb{T}^d))$ that satisfies (CE) (indeed $b^\delta \to b$ strongly and you can pass to the limit in every term of (2.1.1)). Considering (TE), for each fixed δ we have

$$\rho^\delta(t, x) = \rho_0(Y_t^\delta(t, x)),$$

and changing variables as above:

$$\int_{\mathbb{T}^d} |\rho^\delta(t,x)|^q dx = \int_{\mathbb{T}^d} |\rho_0(y)|^q |\det(\nabla_y X^\delta(t,y))| dy. \tag{2.1.4}$$

By the uniform bound on the determinant, given by Proposition 1.1.5, and by the condition on $\mathrm{div}_x(b)^+$, we conclude by weak-$*$ convergence in $L^\infty(0,T;L^q(\mathbb{T}^d))$ as before. If $q=+\infty$, we just need to use weak-$*$ compactness in $L^\infty(I\times\mathbb{T}^d)$ and the definition of weak solution for (TE) in (2.1.2), and no assumption on $\mathrm{div}_x(b)$ is required. We conclude similarly for (CE), where however a condition on $\mathrm{div}_x(b)^-$ is still needed in order to have $\{u^\delta\}$ uniformly bounded in $L^\infty(\mathbb{T}^d)$. Finally, we consider the case $q=1$. We shall prove weak compactness for the sequence $\{\rho^\delta\}_{\delta>0}$ of solutions of (TE) with velocity fields b^δ in $L^1((0,T)\times\mathbb{T}^d)$. To do so, we will use the Dunford-Pettis theorem [34], proving that the sequence is equi-integrable. We recall that a family $\mathcal{F}$ of functions defined on a set $\Omega\subseteq\mathbb{R}^d$ is called equi-integrable if for every $\varepsilon>0$ there exists $\varepsilon'>0$ such that if $A\subseteq\Omega$ with $\mathscr{L}^d(A)\le\varepsilon'$, then

$$\int_A f\,dx \le \varepsilon \quad \forall f\in\mathcal{F}.$$

For every $t\in I$, by the additional assumption on $\mathrm{div}_x(b)^-$ we have that given a set $A\subseteq\mathbb{T}^d$:

$$\mathscr{L}^d(Y_t^\delta(t,A)) \le C\mathscr{L}^d(A) \text{ for some positive constant } C>0, \tag{2.1.5}$$

i.e. the flow of b compresses sets in a uniformly bounded (both in time and with respect to δ) way. This follows easily by the bounds on the determinant we obtained in Proposition 1.1.5 and changing variables in $\mathscr{L}^d(Y_t^\delta(t,A)) = \int_{Y_t^\delta(t,A)} dx$. We have to prove that for every $\varepsilon>0$ there exists $\varepsilon'>0$ such that

$$\mathscr{L}^{d+1}(\Omega)\le\varepsilon' \implies \int_\Omega |\rho^\delta| dxdt \le \varepsilon, \text{ with } \Omega\subset I\times\mathbb{T}^d.$$

It holds that:

$$\int_\Omega |\rho^\delta|(t,x)dxdt = \int_\Omega |\rho_0|(Y_t(t,x))dxdt = \int_{\Omega'_\delta} |\rho_0(y)||\det(\nabla_y X^\delta(t,y))|dydt$$

where $\Omega'_\delta = \{(t,y) \text{ such that } (t,y)=(t,Y_t^\delta(t,x)) \text{ with } (t,x)\in\Omega\}$. In the last equality we used the change of variables $(t,y)=(t,Y_t^\delta(t,x))$ and that

$$dy = \nabla_x Y_t^\delta(t,x)dx \implies dx = [\nabla_x Y_t^\delta(t,x)]^{-1}dy = \nabla_x X^\delta(t,Y_t^\delta(t,x))dy = \nabla_x X^\delta(t,y)dy.$$

We can easily see, by (2.1.5), that

$$\mathscr{L}^{d+1}(\Omega'_\delta) \le C\mathscr{L}^{d+1}(\Omega) \le C\varepsilon' \text{ for every } \delta.$$

Since the determinant is uniformly bounded:

$$\int_\Omega |\rho^\delta|(t,x)dxdt \lesssim \sup_{\substack{A\subset[0,T]\times\mathbb{T}^d\\ \mathscr{L}^{d+1}(A)\le C\varepsilon'}} \int_A |\rho_0(x)|dxdt \le \varepsilon$$

if ε' is chosen small enough, by the absolute continuity of the integral. So $\{\rho^\delta\}$ is weakly relatively compact in $L^1((0,T)\times\mathbb{T}^d)$, and it's easy to see that a weak limit ρ satisfies (2.1.2). As for th continuity equation, the proof follows in exactly the same way. Notice though that we do not need any assumptions on $\mathrm{div}_x(b)^+$ since if $q=1$ after the change of variables the determinants simplify: see (2.1.3). This is not surprising, as we already remarked how the continuity equation (on a set with null boundary) preserves the L^1 norm. □

We now start addressing the problem of uniqueness of solutions. Checking the proof of uniqueness in the smooth classical case, we see that there is a problem in the strategy used for (TE). In Proposition 1.2.1 we used a sort of method of characteristics, which relied on having a pointwise well-defined (backward) flow for the vector field b. This was possible by the classical theory for ODEs, which requires a regularity on b that we do not have here with merely Sobolev vector fields. Namely, we have not (yet) defined a notion of flow for b in this low-regularity setting. As for the uniqueness of solutions of (CE), that strategy seems more promising. In Proposition 1.2.1 we prove uniqueness without relying on the flow of b, but rather on an energy inequality which shows that, if we have some control on $\operatorname{div}_x(b)^-$, then $\int_{\mathbb{T}^d} u(t,x)^2 dx = 0$ for every $t \in I$ if $u_0 = 0$, and thus the solution of (CE) is unique by linearity of the equation. The problem, however, is that at the heart of the argument lies a differentiation in time of $\int_{\mathbb{T}^d} u(t,x)^2 dx$ which, a priori, is not justified in this non-regular setting. Indeed we have only proved existence of a solution $u \in L^\infty(0,T;L^q(\mathbb{T}^d))$, hence we do not have enough regularity in time to differentiate the energy.

The idea of the *renormalization technique*, used in this setting for the first time in [70], is essentially to prove that such a differentiation is possible. We now define the concept of *renormalized solution* and we will later show how, if we have a renormalized solution, we also have uniqueness under suitable assumptions. We will work, following [70] (up to the sign of the velocity b), with the equation:

$$\begin{cases} \partial_t \rho + b \cdot \nabla_x \rho + c\rho = 0 \\ \rho(0,\cdot) = \rho_0 \end{cases}, \tag{DP-L}$$

with $c \in L^1(0,T;L^p(\mathbb{T}^d))$ and looking for solutions $\rho \in L^\infty(0,T;L^q(\mathbb{T}^d))$. Notice that if we can prove the uniqueness of solutions of (DP-L) we automatically get uniqueness for (CE) and (TE). Namely, we just have to choose $c = \operatorname{div}_x(b)$ and $c = 0$ to reduce (DP-L) to (CE) and (TE) respectively. The existence of solutions of (DP-L) for a general $c \in L^1(0,T;L^p(\mathbb{T}^d))$ follows by arguments similar to the ones used in Proposition 2.1.1, see [70, Proposition II.1].

Definition 2.1.1. ***Renormalized solutions***
Let ρ be a solution of (DP-L). *We say that ρ is a renormalized solution of* (DP-L) *if, for every $\beta \in C^1(\mathbb{R})$ with $|\beta'|$ bounded, $\beta(\rho)$ satisfies:*

$$\begin{cases} \partial_t[\beta(\rho)] + b \cdot \nabla_x[\beta(\rho)] + c\rho\beta'(\rho) = 0 \text{ in } (0,T)\times\mathbb{T}^d \\ \beta(\rho(0,x)) = \beta(\rho_0) \end{cases}, \tag{2.1.6}$$

that is:

$$\int_0^T \int_{\mathbb{T}^d} \beta(\rho)[\partial_t\varphi + \operatorname{div}_x(b)\varphi + b\cdot\nabla_x\varphi)] - c\rho\beta'(\rho)\varphi dx dt + \int_{\mathbb{T}^d} \beta(\rho_0)\varphi(0,x)dx = 0 \quad \forall \varphi \in C_c^\infty(I\times\mathbb{T}^d).$$

Remark 2.1.2. ***The heuristic idea of the renormalization technique***
The idea behind the definition of renormalized solutions is, as already mentioned, to translate the validity of the chain rule to a low-regularity setting. Indeed, if ρ solved (DP-L) *in a pointwise sense, then (because of the chain rule) $\beta(\rho)$ would solve* (2.1.6)*:*

$$\begin{aligned} \frac{d}{dt}[\beta(\rho)] + b\cdot\nabla_x[\beta(\rho)] + c\rho\beta'(\rho) &= \beta'(\rho)\partial_t\rho + \beta'(\rho) b\cdot\nabla_x\rho + c\rho\beta'(\rho) \\ &= \beta'(\rho)[\partial_t\rho + b\cdot\nabla_x\rho + c\rho] = 0. \end{aligned}$$

The following result is the crucial property which ultimately enables us to prove uniqueness of solutions.

Theorem 2.1.1. ***Renormalization property***
Let $b \in [L^1(0,T;W^{1,p}(\mathbb{T}^d)) \cap L^\infty([0,T)\times\mathbb{T}^d)]^d$ with $p \geq 1$. Then it has the renormalization property *in $L^\infty(0,T;L^q(\mathbb{T}^d))$ (with $q \in [1,+\infty]$ and such that $1/p + 1/q \leq 1$), that is: every solution of* (DP-L) *in $L^\infty(0,T;L^q(\mathbb{T}^d))$ is a renormalized solution.*

The importance of the above theorem, whose proof we postpone, is apparent once paired with the next result.

Theorem 2.1.2. ***Uniqueness of renormalized solutions***
Let $q \in [1,+\infty]$ and let b be a bounded vector field such that $(\operatorname{div}_x(b) - qc)^+ = \max\{0, \operatorname{div}_x(b) - qc\} \in L^1(0,T;L^\infty(\mathbb{T}^d))$, with $c \in L^1(0,T;L^p(\mathbb{T}^d))$ and $1/p + 1/q \leq 1$. Then renormalized solutions of (DP-L) *in $L^\infty(0,T;L^q(\mathbb{T}^d))$ are unique.*

It is now trivial to prove the following:

Corollary 2.1.1. *Assume both Theorem 2.1.1 and Theorem 2.1.2 hold. Then, solutions of* (DP-L) *are unique.*

Proof. Assume to have two different solutions of (DP-L). By Theorem 2.1.1 they are both renormalized solutions, and by Theorem 2.1.2 renormalized solutions are unique, thus we have a contradiction. □

We will now prove Theorem 2.1.1 and Theorem 2.1.2, starting from the latter.

Proof of Theorem 2.1.2. Let us assume that $1 < q < +\infty$. By the linearity of (DP-L), we just need to prove that any renormalized solution with initial condition $\rho_0 = 0$ is identically 0. Let $\varepsilon > 0$ and consider $\beta_M(x) := \max\{|x|^q, M\}$ for some $M > 0$. Since by assumption ρ is a renormalized solution, $\beta_M(\rho)$ satisfies:

$$\int_0^T \int_{\mathbb{T}^d} |\rho|^q(\partial_t\varphi + \operatorname{div}_x(b\varphi) - \mathbb{1}_{\{|\rho|<M\}}qc\varphi)dxdt = 0 \text{ for every } \varphi \in C_c^\infty(I \times \mathbb{T}^d), \tag{2.1.7}$$

with $\mathbb{1}_A$ being the characteristic function of a set A. Consider the function

$$f(t) := \int_{\mathbb{T}^d} |\rho(t,x)|^q dx.$$

We will have proved the thesis if we show that $f \equiv 0$. Let $t_0 \in (0,T)$ be fixed and consider a sequence of test functions $\varphi_{t_0,\varepsilon}(t,x) = \varphi_{t_0,\varepsilon}(t)$ defined as

$$\varphi_{t_0,\varepsilon}(t) = \begin{cases} 1 & \text{if } t \le t_0 - \varepsilon \\ 1 - \frac{1}{2\varepsilon}(t - t_0 + \varepsilon) & \text{if } t \in (t_0 - \varepsilon, t_0 + \varepsilon) \\ 0 & \text{if } t_0 \ge t_0 + \varepsilon \end{cases}$$

with ε chosen small enough so that $(t_0 - \varepsilon, t_0 + \varepsilon) \subset I$. Using these test functions in (2.1.7), by Fubini's theorem we have:

$$-\fint_{t_0-\varepsilon}^{t_0+\varepsilon} \int_{\mathbb{T}^d} |\rho|^q dxdt = -\int_0^T \left(\int_{\mathbb{T}^d} (\operatorname{div}_x(b) - \mathbb{1}_{\{|\rho|<M\}}qc)|\rho|^q dx \right) \varphi_{t_0,\varepsilon} dt.$$

Letting $\varepsilon \to 0$ and assuming t_0 to be a Lebesgue point for $f(t)$, we have that by Lebesgue differentiation theorem (for the left-hand side) and dominated convergence (for the right-hand side):

$$-\int_{\mathbb{T}^d} |\rho(t_0,x)|^q dx = -\int_0^{t_0} \int_{\mathbb{T}^d} \left(\operatorname{div}_x(b) - \mathbb{1}_{\{|\rho|<M\}}qc\right) |\rho(t,x)|^q dxdt \quad \text{for almost every } t_0 \in (0,T).$$

Thus, in the limit for $M \to +\infty$

$$-\int_{\mathbb{T}^d} |\rho(t_0,x)|^q dx = -\int_0^{t_0} \int_{\mathbb{T}^d} \left(\operatorname{div}_x(b) - qc\right) |\rho(t,x)|^q dxdt \quad \text{for almost every } t_0 \in (0,T),$$

and:

$$\begin{aligned} f(t_0) = \int_{\mathbb{T}^d} |\rho(t_0,x)|^q dx &\le \int_0^{t_0} \|(\operatorname{div}_x(b) - qc)^+\|_{L^\infty(\mathbb{T}^d)} \left(\int_{\mathbb{T}^d} |\rho(t,x)|^q dx \right) dt \\ &= \int_0^{t_0} \|(\operatorname{div}_x(b) - qc)^+\|_{L^\infty(\mathbb{T}^d)} f(t) dt \quad \text{for almost every } t_0 \in (0,T). \end{aligned} \tag{2.1.8}$$

Since $f(0) = 0$ and by the assumptions on $(\operatorname{div}_x(b) - qc)^+$, we conclude using Grönwall's lemma. If $q = +\infty$, we can simply choose $q = 2$ and conclude as above, while if $q = 1$ we can regularise $\beta_M(x) = \max\{|x|, M\}$ (for example considering $\beta_{M,\delta}(x) = \max\{\sqrt{x^2+\delta^2} - \delta, \text{M}\}$), and conclude passing to the limit first in δ and then in M.

□

Remark 2.1.3. ***Time continuity of solutions***
Theorem 2.1.2 actually shows that the unique solution of (CE) *and* (TE) *in* $L^\infty(0,T;L^q(\mathbb{T}^d))$ *also belongs to* $C(0,T;L^q(\mathbb{T}^d))$ *if* $q<+\infty$*. Indeed if we have a renormalized solution it holds* (2.1.8)*, thus* $f(t)=\|\rho(t,\cdot)\|_{L^q(\mathbb{T}^d)}$ *is continuous by Grönwall's lemma, the absolute continuity of the integral on the right-hand side of* (2.1.8) *and the uniform boundedness of* $f(t)$*. See [70, Corollary II.2] for more details.*

Now, we just have to prove Theorem 2.1.1 to conclude.

Proof of Theorem 2.1.1. We regularise by mollification (DP-L). Let η be a standard mollifier (in the spatial variables only) and let $\rho_\varepsilon=\rho*\eta_\varepsilon$, with $\eta_\varepsilon=\frac{1}{\varepsilon^d}\eta\left(\frac{x}{\varepsilon}\right)$. Then, ρ_ε solves

$$\partial_t\rho_\varepsilon+b\cdot\nabla_x\rho_\varepsilon+c\rho_\varepsilon=r_\varepsilon, \tag{2.1.9}$$

with

$$r_\varepsilon:=b\cdot\nabla_x\rho_\varepsilon-\eta_\varepsilon*(b\cdot\nabla_x\rho)+c\rho_\varepsilon-\eta_\varepsilon*(c\rho) \tag{2.1.10}$$

the so-called the *commutator term*[2]. We recall that, in the distributional sense:

$$b\cdot\nabla_x\rho=\mathrm{div}_x(b\rho)-\mathrm{div}_x(b)\rho.$$

We assume for the moment that $r_\varepsilon\to 0$ as $\varepsilon\to 0$ in $L^1((0,T)\times\mathbb{T}^d)$, proving it afterwards in Theorem 2.1.3. Notice that

$$\partial_t\rho_\varepsilon=-\eta_\varepsilon*(b\cdot\nabla_x\rho)-\eta_\varepsilon*(c\rho),$$

where

$$\begin{aligned}\eta_\varepsilon*(b\cdot\nabla_x\rho)(t,x)&=\int_{\mathbb{T}^d}\eta_\varepsilon(x-y)\,\mathrm{div}_x(b(t,y)\rho(t,y))dy-\int_{\mathbb{T}^d}\eta_\varepsilon(x-y)\,\mathrm{div}_x(b(t,y))\rho(t,y)dy\\&=\int_{\mathbb{T}^d}\rho(t,y)b(t,y)\cdot\frac{\nabla_x\eta_\varepsilon(x-y)}{\varepsilon}dy-\eta_\varepsilon*(\mathrm{div}_x(b)\rho)(t,x)=\left[\frac{\nabla_x\eta_\varepsilon}{\varepsilon}*(b\rho)-\eta_\varepsilon*(\mathrm{div}_x(b)\rho)\right](t,x),\end{aligned} \tag{2.1.11}$$

where we used that $\mathrm{div}_x(\eta_\varepsilon b\rho)=\eta_\varepsilon\,\mathrm{div}(b\rho)+\nabla_x\eta_\varepsilon\cdot(b\rho)$. In particular, $\partial_t\rho_\varepsilon\in L^1((0,T)\times\mathbb{T}^d)$, and since ρ_ε is smooth in space, we conclude that ρ_ε is Sobolev in space-time. By the chain rule for Sobolev functions (see [73, Theorem 4.4]), given $\beta\in C^1$ with β' bounded:

$$\partial_t\beta(\rho_\varepsilon)=\beta'(\rho_\varepsilon)\partial_t\rho_\varepsilon \text{ and } \nabla_x\beta(\rho_\varepsilon)=\beta'(\rho_\varepsilon)\nabla_x\rho_\varepsilon \text{ almost everywhere in } I\times\mathbb{T}^d.$$

Multiplying (2.1.9) by $\beta'(\rho_\varepsilon)\in L^\infty((0,T)\times\mathbb{T}^d)$ we then have:

$$\partial_t[\beta(\rho_\varepsilon)]+b\cdot[\nabla_x\beta(\rho_\varepsilon)]+c\rho_\varepsilon\beta'(\rho_\varepsilon)=\beta'(\rho_\varepsilon)r_\varepsilon. \tag{2.1.12}$$

The left-hand side converges, in the sense of distributions, to the left-hand side of (2.1.6) as $\varepsilon\to 0$, while the right-hand side converges to 0 in $L^1((0,T)\times\mathbb{T}^d)$ since β' is bounded and $r_\varepsilon\to 0$ in $L^1((0,T)\times\mathbb{T}^d)$ by assumption.

□

Remark 2.1.4. ***A weaker assumption on*** r_ε
Notice that, in order to prove the above theorem, it is not really necessary that $r_\varepsilon\to 0$ *strongly in* $L^1((0,T)\times\mathbb{T}^d)$ *as* $\varepsilon\to 0$*. What is really needed is that* $\beta'(\rho_\varepsilon)r_\varepsilon\to 0$ *in the sense of distributions: the strong convergence of the commutator is a sufficient but not necessary condition.*

We are now left to prove that $r_\varepsilon\to 0$ in $L^1((0,T)\times\mathbb{T}^d)$. Before doing that, we prove a technical Lemma which is ultimately the main technical point of the whole argument (we will follow the notation of [70]).

[2]This name comes from the fact that r_ε estimates the error produced when commuting the operator $\mathcal{L}_{b,c}(\rho)=b\cdot\nabla_x\rho+c\rho$ with the mollification.

Lemma 2.1.1. ***Crucial technical lemma***
Let $B \in [L^1(0,T;W^{1,p}(\mathbb{T}^d))]^d$ and $w \in L^\infty(0,T;L^q(\mathbb{T}^d))$ with $1/p+1/q \le 1$, and let η_ε be a standard mollifier (in spatial variables only). Then:

$$(B \cdot \nabla_x w) * \eta_\varepsilon - B \cdot \nabla_x(\eta_\varepsilon * w) \to 0 \text{ in } L^1((0,T)\times\mathbb{T}^d).$$

Proof. As we saw in (2.1.11), $\eta_\varepsilon * (B\cdot\nabla_x w) = \frac{\nabla_x \eta_\varepsilon}{\varepsilon} * (Bw) - \eta_\varepsilon * (\mathrm{div}_x(B)w)$. So:

$$\begin{aligned}&[B\cdot\nabla_x(\eta_\varepsilon * w) - (B\cdot\nabla_x w) * \eta_\varepsilon](t,x) \\ &= \eta_\varepsilon * (\mathrm{div}_x(B)w)(t,x) + \int_{B_\varepsilon(x)} w(t,y)\frac{B(t,x)-B(t,y)}{\varepsilon}\cdot\nabla_x\eta_\varepsilon(x-y)dy,\end{aligned} \tag{2.1.13}$$

where we only integrate in a ball $B_\varepsilon(x)$ of radius ε centered in x since $\mathrm{Supp}(\eta_\varepsilon) \subseteq B_\varepsilon(0)$. We focus on the second addendum. By Hölder's inequality:

$$\left|\int_{B_\varepsilon(x)} w(t,y)\frac{B(t,x)-B(t,y)}{\varepsilon}\cdot\nabla_x\eta_\varepsilon(x-y)dy\right| \lesssim \|w(t,\cdot)\|_{L^q(\mathbb{T}^d)}\|f_{x,\varepsilon}\|_{L^p(B_\varepsilon(x))},$$

with

$$f_{x,\varepsilon}(y) = \frac{B(t,x)-B(t,y)}{\varepsilon}.$$

Let us show that $\|f_{x,\varepsilon}(y)\|_{L^p(B_\varepsilon(x))}$ is uniformly bounded as $\varepsilon \to 0$, so that we can apply Lebesgue's dominated convergence. Indeed:

$$\begin{aligned}\|f_{x,\varepsilon}\|_{L^p(B_\varepsilon)} &= \left(\int_{B_\varepsilon(x)}\left|\frac{B(t,x)-B(t,y)}{\varepsilon}\right|^p dy\right)^{1/p} \lesssim \left(\int_{B_\varepsilon(x)}\left|\int_0^1 |\nabla_x B(t,x+s(y-x))|ds\right|^p dy\right)^{1/p} \\ &\le \left(\int_{B_\varepsilon(x)}\int_0^1 |\nabla_x B(t,x+s(y-x))|^p ds dy\right)^{1/p} \le \|\nabla_x B\|_{L^p(\mathbb{T}^d)}.\end{aligned}$$

By looking at (2.1.13), we just need to prove that

$$\int_{B_\varepsilon} w(t,y)\frac{B(t,x)-B(t,y)}{\varepsilon}\cdot\nabla_x\eta_\varepsilon(x-y)dy \to -[w\,\mathrm{div}_x(B)](t,x) \quad \text{as } \varepsilon \to 0$$

to conclude. Changing variables in the integral with $z=(x-y)/\varepsilon$ yields:

$$\int_{B_1}\omega(t,x-\varepsilon z)\frac{B(t,x)-B(t,x-\varepsilon z)}{\varepsilon}\nabla_z\eta(z)dz \to \omega(t,x)\int_{B_1}\nabla_x B(t,x)z\nabla_z\eta(z)dz$$

as $\varepsilon \to 0$, where

$$\nabla_x B(t,x)z\nabla_z\eta(z) := \sum_{i,j=1}^d \frac{\partial B^i}{\partial x_j}(t,x)z^j\frac{\partial\eta}{\partial z_i}(z).$$

So

$$\int_{B_1}\nabla B(t,x)z\nabla_z\eta(z)dz = \sum_{i,j=1}^d\frac{\partial B^i}{\partial x_j}(t,x)\left(\int_{B_1} z^j\frac{\partial\eta}{\partial z_i}(z)dz\right) = -\mathrm{div}_x(B)(t,x),$$

since one can check easily that considering η a standard mollifier

$$\int_{B_1} z^j\frac{\partial\eta}{\partial z_i}dz = -\delta_{ij},$$

with δ_{ij} the Kronecker delta. We then have:

$$\int_{B_\varepsilon} w(t,y)\frac{B(t,x)-B(t,y)}{\varepsilon}\cdot\nabla_x\eta_\varepsilon(x-y)dy \to -w(t,x)\operatorname{div}_x(B)(t,x) \quad \text{as } \varepsilon\to 0,$$

and we also have that the first term in (2.1.13) tends to $w(t,x)\operatorname{div}_x(B)(t,x)$ as $\varepsilon\to 0$. So, for almost every (t,x), it holds that

$$[B\cdot\nabla_x(\eta_\varepsilon * w)-(B\cdot\nabla_x w)*\eta_\varepsilon](t,x)\to 0 \text{ in a pointwise sense.}$$

Since by the previous arguments it is easy to see that $B\cdot\nabla_x(\eta_\varepsilon * w)-(B\cdot\nabla_x w)*\eta_\varepsilon$ is uniformly bounded (with respect to ε), we have the thesis by dominated convergence. □

We are now ready to prove the strong convergence of the commutator term to 0.

Theorem 2.1.3. ***Strong convergence of the commutator***
Let r_ε as in (2.1.10). *If $b\in[L^1(0,T;W^{1,p}(\mathbb{T}^d))\cap L^\infty((0,T)\times\mathbb{T}^d)]^d$ with $p\ge 1$, $c\in L^1(0,T;L^p(\mathbb{T}^d))$ and we assume $\rho\in L^\infty(0,T;L^q(\mathbb{T}^d))$ with $1/p+1/q\le 1$, then*

$$r_\varepsilon\to 0 \textit{ in the strong } L^1((0,T)\times\mathbb{T}^d) \textit{ norm as } \varepsilon\to 0.$$

Proof. We have that $r_\varepsilon = r^1_\varepsilon + r^2_\varepsilon$, with

$$r^1_\varepsilon = b\cdot\nabla_x\rho_\varepsilon - \eta_\varepsilon * (b\cdot\nabla_x\rho)$$

and

$$r^2_\varepsilon = c\rho_\varepsilon - \eta_\varepsilon * (c\rho).$$

Concerning r^2_ε, it is easy to see that $r^2_\varepsilon\to 0$ in $L^1((0,T)\times\mathbb{T}^d)$ since

$$c\rho_\varepsilon\to c\rho \quad\text{and}\quad \eta_\varepsilon*(c\rho)\to c\rho \quad \text{in } L^1((0,T)\times\mathbb{T}^d) \text{ as } \varepsilon\to 0$$

by standard properties of the mollifiers. The same holds for r^1_ε, by Lemma 2.1.1. □

To summarise the results obtained so far, we state a theorem combining the existence and uniqueness results for (CE) and (TE).

Theorem 2.1.4. ***Existence and uniqueness of solutions of* (CE), (TE) *with Sobolev vector fields, [70]***
Let $b\in[L^1(0,T;W^{1,p}(\mathbb{T}^d))\cap L^\infty((0,T)\times\mathbb{T}^d)]^d$ with $p\ge 1$ and let $q\in[1,+\infty]$ such that $1/p+1/q\le 1$. Then:

- *If $\operatorname{div}_x(b)^-\in L^1(0,T;L^\infty(\mathbb{T}^d))$,* (CE) *admits a unique solution in $L^\infty(0,T;L^q(\mathbb{T}^d))$;*
- *If $q\in(1,+\infty)$ and $\operatorname{div}_x(b)^+\in L^1(0,T;L^\infty(\mathbb{T}^d))$,* (TE) *admits a unique solution in $L^\infty(0,T;L^q(\mathbb{T}^d))$. If $q=+\infty$ no condition on the divergence is necessary to have existence, while for $q=1$ existence holds under the stronger assumption $\operatorname{div}_x(b)\in L^1(0,T;L^\infty(\mathbb{T}^d))$.*

Well-posedness has also been proved for BV vector fields and bounded densities in [7], where the author used Alberti's rank one lemma and local anisotropic convolution kernels inspired by [29, 50] to show that $r_\varepsilon\to 0$ in $L^1((0,T)\times\mathbb{T}^d)$, proving the analogous of Theorem 2.1.1 in the BV setting. Uniqueness then follows by Theorem 2.1.2, where we recall no regularity on b was needed, and Corollary 2.1.1.

Theorem 2.1.5. ***Existence and uniqueness with BV vector fields, [7]***
Let $b\in[L^1(0,T;BV(\mathbb{T}^d))\cap L^\infty((0,T)\times\mathbb{T}^d)]^d$ with $\operatorname{div}_x(b)^-\in L^1(0,T;L^\infty(\mathbb{T}^d))$ and assume $\operatorname{div}_x(b)$ is an absolutely continuous measure with respect to $\mathscr{L}^d$. Then b has the renormalization property, i.e. every solution ρ of (DP-L) *with vector field b and $c\in L^1((0,T)\times\mathbb{T}^d)$ is a renormalized solution. In particular,* (CE) *and* (TE) *both have a unique solution in $L^\infty((0,T)\times\mathbb{T}^d)$.*

Another extension has been proved in [21] for *nearly incompressible vector fields*[3]. Before stating the main result of [21], we give the definition of a nearly incompressible vector field.

Definition 2.1.2. ***Nearly incompressible vector fields, [21]***
A locally integrable vector field $b : \mathbb{R}_0^+ \times \mathbb{T}^d \to \mathbb{R}^d$ is called nearly incompressible if there exists a function $\rho : \mathbb{R}_0^+ \times \mathbb{T}^d \to \mathbb{R}^+$ (called the density of the vector field) and a constant $C > 0$ such that $C^{-1} \leq \rho(t,x) \leq C$ for $\mathscr{L}^{d+1}$-almost every $(t,x) \in \mathbb{R}_0^+ \times \mathbb{T}^d$ and

$$\partial_t \rho + \operatorname{div}_x(\rho b) = 0 \text{ in the sense of distributions on } \mathbb{R}_0^+ \times \mathbb{T}^d.$$

Notice that, unlike in the previous cases where absolute continuity of $\operatorname{div}_x(b)$ was assumed, here it is not. In particular, Theorem 2.1.2 (which guarantees uniqueness of solutions for (CE) and (TE) in the previous cases) cannot be applied, and actually (TE) does not even make sense distributionally since the singular part of $\operatorname{div}_x(b)$ (used to define the distribution $b \cdot \nabla_x \rho$, see (2.1.2)) might be non-zero. For this reason, we need another definition of solution and of the renormalized property for nearly incompressible vector fields.

Definition 2.1.3. ***ρ-weak solution***
Let $I = [0,T)$ be a time interval and let b be a bounded nearly incompressible vector field with density ρ. We say that a function $u \in L^\infty((0,T) \times \mathbb{T}^d)$ is a ρ-weak solution to the transport equation $\partial_t u + b \cdot \nabla_x u = 0$ if

$$\partial_t(\rho u) + \operatorname{div}_x(\rho u b) = 0 \text{ in the sense of distributions in } I \times \mathbb{T}^d. \tag{ρ-TE}$$

Definition 2.1.4. ***Renormalized property for nearly incompressible vector fields***
Let $I = [0,T)$ be a time interval. We say that a bounded nearly incompressible vector field $b : I \times \mathbb{T}^d \to \mathbb{R}^d$ has the renormalization property if, given its density ρ as in Definition 2.1.2, every solution $u \in L^\infty((0,T) \times \mathbb{R}^d)$ of

$$\partial_t(\rho u) + \operatorname{div}_x(\rho u b) = 0$$

satisfies

$$\partial_t(\rho \beta(u)) + \operatorname{div}_x(\rho \beta(u) b) = 0$$

for every $\beta \in C^1(\mathbb{R})$ with β' bounded.

The main result proved in [21] is then:

Theorem 2.1.6. ***Renormalization property for nearly incompressible vector fields***
Any nearly incompressible vector field $b \in [L^1([0,T), BV_{loc}(\mathbb{R}^d)) \cap L^\infty((0,T) \times \mathbb{T}^d)]^d$ has the renormalization property.

The existence and uniqueness of solutions of (ρ-TE) when the vector field is nearly incompressible and assuming the renormalization property holds is proved in [61, Proposition 3.13]. We remark that this result has answered in the positive a conjecture proposed by Bressan [32] concerning stability of flows (this implication is not trivial and it has been proved in [11]):

Theorem 2.1.7. ***Bressan's compactness conjecture, [32]***
Let $b_n : \mathbb{R} \times \mathbb{T}^d \to \mathbb{R}^d$, $n \in \mathbb{N}$, be a sequence of smooth vector fields. Denote by Φ_n the flow generated by b_n, i.e. Φ_n solves

$$\begin{cases} \frac{d}{dt}\Phi_n(t,x) = b_n(t, \Phi_n(t,x)) \\ \Phi(0,x) = x. \end{cases}$$

Assume that $\|b_n\|_\infty + \|\nabla_{t,x} b_n\|_{L^1}$ is uniformly bounded and there exists a constant $C > 0$ such that

$$C^{-1} \leq \det(\nabla_x \Phi_n(t,x)) \leq C \text{ for all } (t,x) \in \mathbb{R} \times \mathbb{T}^d \text{ and for all } n \in \mathbb{N}.$$

Then the sequence $\{\Phi_n\}_{n\in\mathbb{N}}$ is strongly precompact in L^1.

[3] We cite the results in [7] and [21] considering the vector field defined in $\mathbb{T}^d$, while in [7], [21] they are defined in the whole $\mathbb{R}^d$. This is just to keep the convention of working in $\mathbb{T}^d$.

To summarise the results, we explicitly state a theorem.

Theorem 2.1.8. ***Existence and uniqueness for nearly incompressible vector fields, [21, 61]***
Let $b \in [L^1([0,T), BV_{loc}(\mathbb{R}^d)) \cap L^\infty((0,T)\times\mathbb{T}^d)]^d$ and nearly incompressible. Then (ρ-TE) *admits a unique solution in $L^\infty((0,T)\times\mathbb{T}^d)$.*

We conclude this section proving the stability of solutions of (CE), (TE) when approximating the initial data and the vector field b. As we will see in the proof, we will exploit the fact that solutions are unique by having a unique limit for the approximating functions. For simplicity we only consider bounded initial data and we prove the stability in the class $L^\infty((0,T)\times\mathbb{R}^d)$, but the result can be generalised to $L^\infty(0,T;L^q(\mathbb{T}^d))$ for other values of q provided that the vector field b has the right $L^1_t W^{1,p}_x$ regularity.

Theorem 2.1.9. ***Stability of solutions of*** (CE)***,*** (TE)
Let $\{b_n\}_{n\in\mathbb{N}}$ be a sequence of smooth bounded vector fields such that

$$b_n \to b \quad \text{in } [L^1((0,T)\times\mathbb{T}^d)]^d, \quad \text{with } b \in [L^1(0,T;W^{1,p}\cap L^\infty(\mathbb{T}^d))]^d$$

and

$$\operatorname{div}(b_n) \to \operatorname{div}(b) \quad \text{in } L^1((0,T)\times\mathbb{T}^d), \quad \text{with } \operatorname{div}_x(b) \in L^1(0,T;L^\infty(\mathbb{T}^d)).$$

Let $\{\rho_0^n\}_{n=1}^\infty$, $\{u_0^n\}_{n\in\mathbb{N}}$ be a sequence of equibounded functions converging in $L^1(\mathbb{T}^d)$ to the initial data ρ_0 of (TE) *and u_0 of* (CE)*, respectively. Then the solutions ρ_n of* (TE) *with initial datum ρ_0^n and driving vector fields b_n converge in $L^2((0,T)\times\mathbb{T}^d)$ to the unique solution of* (TE) *with driving vector field b and initial datum ρ_0. The same holds for* (CE) *assuming also that the sequence $\{\|\operatorname{div}_x(b_n)^-\|_{L^1(0,T;L^\infty(\mathbb{T}^d))}\}_{n\in\mathbb{N}}$ is equibounded.*

Proof. We start considering (TE). For each n, by the classical theory there exists a unique solution ρ_n for (TE) with vector field b_n and initial datum ρ_0^n, which is given by the explicit formula (1.2.3). From (1.2.3) it immediately follows that the sequence ρ_n is equibounded and thus (up to a subsequence we will not rename) we have weak-$*$ convergence to a function ρ which (thanks to the uniqueness results proved previously) is the unique solution of (TE) with vector field b and initial datum ρ_0. Indeed for every $\varphi \in C_c^\infty(I\times\mathbb{T}^d)$ if we pass to the limit for $n \to +\infty$ in

$$\int_0^T\int_{\mathbb{T}^d} \rho_n(\partial_t\varphi + \operatorname{div}_x(b_n)\varphi + b_n\cdot\nabla_x\varphi)dxdt + \int_{\mathbb{T}^d}\varphi(0,x)\rho_0^n(x)dx = 0$$

we easily obtain that ρ solves

$$\int_0^T\int_{\mathbb{T}^d} \rho(\partial_t\varphi + \operatorname{div}_x(b)\varphi + b\cdot\nabla_x\varphi)dxdt + \int_{\mathbb{T}^d}\varphi(0,x)\rho_0(x)dx = 0.$$

If we repeat the argument with initial datum $(\rho_0^n)^2$, we have that the solutions ρ_n^2 must converge to the unique solution of (TE) with initial datum $(\rho_0)^2$ and velocity field b. But by the renormalization property this solution is ρ^2. So, we have weak-$*$ convergence of ρ_n^2, which implies (by choosing a test function φ_ε such that $\varphi_\varepsilon \equiv 1$ in $[0,T-\varepsilon)\times\mathbb{T}^d$ and passing to the limit as $\varepsilon \to 0$)

$$\int_0^T\int_{\mathbb{T}^d}\rho_n^2 dxdt \to \int_0^T\int_{\mathbb{T}^d}\rho^2 dxdt \quad \text{as } n\to+\infty.$$

Since weak-$*$ convergence in $L^\infty((0,T)\times\mathbb{T}^d)$ implies weak convergence in $L^2((0,T)\times\mathbb{T}^d)$ we have both weak convergence in $L^2((0,T)\times\mathbb{T}^d)$ of ρ_n to ρ and $\|\rho_n\|_{L^2((0,T)\times\mathbb{T}^d)} \to \|\rho\|_{L^2((0,T)\times\mathbb{T}^d)}$. This implies, by a well-known result in functional analysis, that $\rho_n \to \rho$ in $L^2((0,T)\times\mathbb{T}^d)$. The case for (CE) follows in the same way. The boundedness condition on $\{\|\operatorname{div}_x(b_n)^-\|_{L^1(0,T;L^\infty(\mathbb{T}^d))}\}_{n=1}^\infty$ is needed in order to have that the solutions u_n of (CE) with respect to the vector fields b_n are equibounded in $L^\infty((0,T)\times\mathbb{T}^d)$, using the representation formula (1.2.4).

□

Remark 2.1.5. ***Necessity of strong convergence of*** $\mathbf{div_x(b_n)}$
Notice that in the proof of Theorem 2.1.9 the weak-$*$ convergence of the solutions u_n to a solution u of (CE) *was guaranteed even if we did not assume the strong convergence of $\operatorname{div}_x(b_n)$. Indeed, contrary to*

(2.1.2), *the integral formulation of the continuity equation* (2.1.1) *does not contain any divergence of the vector field and weak-$*$ convergence is sufficient to pass to the limit. However, as noted in Remark 2.1.6, strong convergence fails since the integral equation* (2.1.6) *solved by $\beta(u_n) = u_n^2$ can be rewritten as $\partial_t(u_n^2) + \mathrm{div}_x(b_n u_n^2) + \mathrm{div}_x(b_n)u_n^2 = 0$, thus the product $\mathrm{div}_x(b_n)u_n$ needs to converge and since we only have that u_n converges weakly-$*$, we need strong convergence of $\mathrm{div}_x(b_n)$. For the same reason (at least, in principle) one would expect that the strong convergence of $\mathrm{div}_x(b_n)$ is required for having convergence (even weak!) of ρ_n to ρ, since in* (2.1.2) *there is a product $\mathrm{div}(b_n)\rho_n$ which needs to converge to $\mathrm{div}_x(b)\rho$. As we will see in Proposition 4.2.3 this is false, and actually the weak convergence of $\{\mathrm{div}_x(b_n)\}$ in $L^1(0,T;L^\infty(\mathbb{T}^d)$ is enough to prove strong convergence of ρ_n to ρ. We postpone this proof since we need to develop the theory of regular Lagrangian flows in Section 2.2 and some key estimates in Chapter 4.*

Remark 2.1.6. ***A counterexample to strong convergence of solutions***
It is shown via a counterexample in [62] (see also [145, Example 3.2]) that, in general, if we have a sequence of vector fields $\{b_n\}_{n\in\mathbb{N}}$ converging strongly to b but whose divergences converge only weakly to $\mathrm{div}_x(b)$, then we cannot conclude that the solutions u_n of (CE) *with respect to b_n converge to the solution u of* (CE) *relative to b in strong L^1 norm.*

Remark 2.1.7. ***The 1D and 2D cases***
It is worth mentioning that the cases $d = 1$ and $d = 2$ are somewhat special. If $d = 1$, uniqueness of solutions in the class $L^1_{loc}([0,T) \times \mathbb{R})$ has been proved in [88] for bounded nearly incompressible vector fields (in [21], where the general case is treated, uniqueness was proved in the L^∞ class). If $d = 2$, in [5] the authors characterise the existence and uniqueness of a bounded solution for (CE) *in terms of a property on the Lipschitz potential of the vector field.*

2.2 The ODE theory: Ambrosio's superposition principle

2.2.1 The general theory of regular Lagrangian flows

Before presenting the theory for ODEs relative to non-regular velocity fields, we give some relevant definitions and results with smooth vector fields to fix the ideas [4]. As we briefly mentioned in Section 1.2, the continuity equation is well-posed even in more general settings; namely, we can formulate the continuity equation as a *measure -valued PDE*.

Definition 2.2.1. ***Measure-valued continuity equation***
Let $\mathcal{M}(\mathbb{T}^d)$ be the space of finite signed measures on $\mathbb{T}^d$. Let $\bar{\mu} \in \mathcal{M}(\mathbb{T}^d)$ and b be a bounded vector field. We say that the family of locally finite signed measures $\{\mu_t\}_{t=0}^T$ solves the (measure-valued) continuity equation in $I = [0,T)$ if it solves

$$\begin{cases} \partial_t \mu_t + \mathrm{div}_x(b\mu_t) = 0 & \text{for } t \in (0,T) \\ \mu_0 = \bar{\mu} \end{cases}, \qquad (\mu\text{-CE})$$

that is if

$$\int_0^T \int_{\mathbb{T}^d} (\partial_t \varphi + b \cdot \nabla_x \varphi) d\mu_t dt + \int_{\mathbb{T}^d} \varphi(0,x)\bar{\mu} = 0 \quad \text{for every } \varphi \in C_c^\infty(I \times \mathbb{T}^d).$$

Remark 2.2.1. ***The choice of the class where to prove well-posedness***
We notice that in order to prove existence and uniqueness of solutions of (CE), *it is crucial to specify both a subset of $\mathcal{M}(\mathbb{T}^d)$ of admissible solutions and assumptions on b and on $\bar{\mu}$ such that we can prove both existence and uniqueness in the chosen subset. Of course, there is a trade-off: if the chosen subset is bigger, proving existence is easier. But proving the uniqueness may be harder. For instance, in Proposition 1.2.1 we proved existence and uniqueness of solutions in the class $C^\infty([0,T) \times \mathbb{T}^d)$ under suitable assumptions on the vector field b and on the initial datum u_0 . In Theorem 2.1.4 we asked for less regularity on b and on the initial datum, but we had to consider a bigger subset of $\mathcal{M}(\mathbb{T}^d)$.*

[4] Also in this section we will state the results considering the domain to be $\mathbb{T}^d$ and bounded vector fields for simplicity. We refer the reader to [10] for the unbounded case in $\mathbb{R}^d$.

We now prove an existence result analogous to Proposition 1.2.1.

Proposition 2.2.1.
Consider b a smooth vector field. Then, for any initial datum $\bar{\mu} \in \mathcal{M}(\mathbb{T}^d)$, a solution of ($\mu$-CE) is given by

$$\mu_t = X(t, \cdot)_{\#}\bar{\mu},$$

where X is the flow of the vector field b and $X(t,\cdot)_{\#}\bar{\mu}$ denotes a push-forward via the map $X(t,\cdot) : \mathbb{T}^d \to \mathbb{T}^d$ of the measure $\bar{\mu}$. In particular, if $\bar{\mu} = u_0 \mathscr{L}^d$ with $u_0 \in L^\infty(\mathbb{T}^d)$, there is an explicit formula for the solution, namely:

$$u(t,x) = \frac{u_0(Y_t(t,x))}{\det(\nabla_x X(t, Y_t(t,x)))},$$

with Y_t the backward flow of b.

Proof. We simply check that the proposed measure μ_t is a solution in the sense of Definition 2.2.1. By the push-forward formula:

$$\int_0^T \int_{\mathbb{R}^d} (\partial_t \varphi + b \cdot \nabla_x \varphi) d(X(t,\cdot)_{\#}\bar{\mu}) dt = \int_0^T \int_{\mathbb{R}^d} (\partial_t \varphi(t, X(t,x)) + b(t, X(t,x)) \cdot \nabla_x \varphi(t, X(t,x))) d\bar{\mu} dt$$
$$= \int_0^T \int_{\mathbb{R}^d} \frac{d}{dt} \left[\varphi(t, X(t,x))\right] d\bar{\mu} dt.$$

We can easily conclude by applying Fubini theorem:

$$\int_0^T \int_{\mathbb{R}^d} \frac{d}{dt} \left[\varphi(t, X(t,x))\right] d\bar{\mu} dt = \int_{\mathbb{R}^d} \int_0^T \frac{d}{dt} \left[\varphi(t, X(t,x))\right] d\bar{\mu} dt$$
$$= \int_{\mathbb{R}^d} \underbrace{\varphi(T, X(T,x))}_{=0} d\bar{\mu} - \int_{\mathbb{R}^d} \varphi(0, X(0,x)) d\bar{\mu} = - \int_{\mathbb{R}^d} \varphi(0,x) d\bar{\mu}.$$

The validity of the representation formula if $\bar{\mu} = u_0 \mathscr{L}^d \in L^\infty(\mathbb{T}^d)$ has already been proved in Proposition 1.2.1. □

A general theory relating the Eulerian/PDE point of view to the Lagrangian/ODE one in a non-classical setting has been developed in [10, 57], which we refer the reader to for more details on this topic. The basic idea is similar to what we considered in Proposition 2.2.1: the initial measure $\bar{\mu}$ is transported via (a generalised, weaker notion of) the flow X and the solution of (μ-CE) is simply the push-forward of the initial datum via such a flow. For the sake of readability, we recall here the equation for the flow X of b:

$$\begin{cases} \frac{d}{dt} X(t,x) = b(t, X(t,x)) & \text{if } t \in (0,T) \\ X(0,x) = x \end{cases} . \tag{ODE}$$

We remark that, in order to give a meaning to (ODE), even in an almost everywhere sense, we assume b to be defined pointwise in $[0,T) \times \mathbb{T}^d$. Also, in view of the above discussion, it is convenient to define the notion of *Lagrangian solution*:

Definition 2.2.2. *Lagrangian solution*
Consider b a bounded measurable vector field in $I \times \mathbb{T}^d$, and let $\bar{\mu} \in \Omega \subset \mathcal{M}(\mathbb{T}^d)$ be a finite signed measure. We say that μ_t is a Lagrangian solution of (μ-CE) in the set Ω if for every t

$$\mu_t = X(t,\cdot)_{\#}\bar{\mu},$$

where $X(\cdot, x)$ solves (ODE) for $\bar{\mu}$-a.e. $x \in \mathbb{T}^d$ and it holds that $X(t,\cdot)_{\#}\bar{\mu} \in \Omega$ for every $t \in I$.

Remark 2.2.2. *Lagrangian $\implies$ renormalized*
It is not difficult to prove that, if $\Omega = L^\infty(0,T; L^q(\mathbb{T}^d))$, Lagrangian solutions are also renormalized solutions [51].

The first result we present is a general criterion relating uniqueness of solutions of (μ-CE) and (ODE):

Proposition 2.2.2. ***[10]***
Let $A \subset \mathbb{R}^d$ be a Borel set. The following two properties are equivalent:

1. *Solutions of* (ODE) *are unique for every $x \in A$.*
2. *Non-negative measure-valued solutions of* (μ-CE) *are unique for any $\bar{\mu}$ concentrated in A, i.e. such that $\bar{\mu}(A^c) = 0$.*

The implication $2. \implies 1.$ is rather easy: just choose $\bar{\mu} = \delta_x$, where δ_x is a Dirac delta centered in x, and notice that different solutions X_1, X_2 of (ODE) induce different solutions of (μ-CE): $\mu_1(t) = \delta_{X_1(t,x)}$ and $\mu_2(t) = \delta_{X_2(t,x)}$. The converse implication requires a deep result known as *superposition principle* (first established in [13, Theorem 8.2.1] , see also [8, Theorem 3.2] and [10, 57]), which states that every *positive* solution of (μ-CE) can be decomposed (as a measure) along characteristic curves solving (ODE), for every $x \in A$. Namely, it exploits the representation

$$\int_{\mathbb{T}^d} \varphi d\mu_t = \int_{\mathbb{T}^d} \left(\int_{\Gamma_T} \varphi(\gamma(t)) d\eta_x(\gamma) \right) d\bar{\mu}(x),$$

obtained by disintegrating the measure (see [12, Theorem 2.28]) μ_t with respect to $\bar{\mu}$, where η_x are probability measures concentrated on absolutely continuous integral solutions of (ODE) with initial point x and $\Gamma_T = AC(0,T;\mathbb{R}^d)$ the set of absolutely continuous curves from $[0,T)$ to $\mathbb{R}^d$. Before stating the superposition principle and discussing its implications, we give the definition of *superposition solution*:

Definition 2.2.3. ***Superposition solution***
Let $\eta \in \mathcal{M}_+(\mathbb{T}^d \times \Gamma_T)$ be a positive measure concentrated on the set of pairs (x,γ) such that γ is an absolutely continuous integral solution of (ODE) *with $\gamma(0) = x$. We define the superposition solution μ_t^η as:*

$$\langle \mu_t^\eta, \varphi \rangle := \int_{\mathbb{T}^d \times \Gamma_T} \varphi(\gamma(t)) d\eta(x,\gamma) \text{ for every } \varphi \in C_b(\mathbb{T}^d).$$

By disintegrating the measure μ_t^η as $\mu_t^\eta = \eta_x(\gamma) \otimes \bar{\mu}(x)$ we can rewrite the definition as

$$\langle \mu_t^\eta, \varphi \rangle = \int_{\mathbb{T}^d} \left(\int_{\Gamma_T} \varphi(\gamma(t)) d\eta_x(\gamma) \right) d\bar{\mu}(x), \text{ for every } \varphi \in C_b(\mathbb{T}^d),$$

where η_x are probability measures on Γ_T.

Remark 2.2.3. ***Recovering of the push-forward formula***
Notice that if $\eta_x = \delta_{X(\cdot,x)}$, i.e. if there is only one absolutely continuous solution to (ODE) *with initial datum $\gamma(0) = x$, we have:*

$$\langle \mu_t^\eta, \varphi \rangle = \int_{\mathbb{T}^d} \varphi(X(t,x)) d\bar{\mu}(x) = X(t,\cdot)_\# \bar{\mu},$$

recovering the push-forward representation of the solution of Proposition 2.2.1.

We are now ready to state the superposition principle, with which it is possible to prove the implication $1. \implies 2.$ in Proposition 2.2.2.

Theorem 2.2.1. ***Superposition principle, [13, 10]***
Let $\mu_t \in \mathcal{M}_+(\mathbb{T}^d)$ be a positive solution of (μ-CE) *with b a bounded vector field. Then μ_t is a superposition solution, i.e. there exists a measure $\eta \in \mathcal{M}_+(\mathbb{T}^d \times \Gamma_T)$ such that $\mu_t = \mu_t^\eta$.*

We remark that Proposition 2.2.2 may seem powerful, but it is in fact quite limited: recall that when b has merely Sobolev spatial regularity, in order to have uniqueness for (μ-CE) we had to restrict to measures $\mu \in L^\infty(0,T;L^q(\mathbb{T}^d))$ (i.e., with a little abuse of notation, to measures which can be represented as $\mu_t = u(t,\cdot)\mathscr{L}^d$ with $u \in L^\infty(0,T;L^q(\mathbb{T}^d))$), which is much more restrictive than considering all non-negative measures as in Proposition 2.2.2. On the other hand, Proposition 2.2.2 states that uniqueness in the whole class of non-negative measures gives a very strong pointwise uniqueness for the flows. For this reason, it is natural to introduce a weaker notion of flow for which uniqueness, assuming (μ-CE) has unique solutions in the smaller class $L^\infty(0,T;L^q(\mathbb{T}^d))$, may still hold. This is the so-called *regular Lagrangian flow*: besides requiring that it solves (ODE) for almost every $x \in \mathbb{T}^d$, we also impose that it does not "concentrate mass". Before giving the actual definition, it is convenient to explain the idea behind it. Since for (μ-CE) we have only proved uniqueness in a subclass of measures (namely absolutely continuous measures with respect to $\mathscr{L}^d$ and with $L^\infty_t L^q_x$ integrability), we want that the flow X which "transports" the initial measure $\bar{\mu}$ preserves the absolute continuity of the measure and the belonging to the $L^\infty_t L^q_x$ class: i.e. $X(t,\cdot)_\#\bar{\mu} \ll \mathscr{L}^d$ for every $t \in [0,T)$ and $\sup_{t\in[0,T]} \|X(t,\cdot)_\#\bar{\mu}\|_{L^q(\mathbb{T}^d)} < +\infty$. It is possible to use the superposition principle not only to prove that $1. \implies 2.$ in the most general (and, as we saw, quite useless) setting of Proposition 2.2.2, but also to prove the implication $2. \implies 1.$ in the subset $L^\infty(0,T;L^q(\mathbb{T}^d))$ (for which we have uniqueness) if we only consider solutions of (ODE) which are regular Lagrangian flows. In particular, we will see how well-posedness results for (2.2.1) in the class $L^\infty((0,T)\times\mathbb{T}^d)$ give well-posedness at the ODE level for regular Lagrangian flows.

Definition 2.2.4. ***Regular Lagrangian flow***
Let $b : [0,T)\times\mathbb{T}^d \to \mathbb{R}^d$ be a bounded vector field. We say that a map $X : [0,T)\times\mathbb{T}^d \to \mathbb{R}^d$ is a regular Lagrangian flow with respect to the vector field b if:

1. *For $\mathscr{L}^d$-a.e. $x \in \mathbb{T}^d$ the map $t \mapsto X(t,x)$ is an absolutely continuous integral solution of $\dot{\gamma}(t) = b(t,\gamma(t))$ for $t \in [0,T)$, with initial condition $\gamma(0) = x$;*
2. *there exists a constant C independent of t such that*

$$X(t,\cdot)_\#\mathscr{L}^d \le C\mathscr{L}^d \text{ as measures.}$$

The constant C will be called the compressibility constant *of X.*

For the sake of completeness, and since we will use it in the following, we also define the *backward* regular Lagrangian flow:

Definition 2.2.5. ***Backward regular Lagrangian flow***
Let $b : [0,T)\times\mathbb{T}^d \to \mathbb{R}^d$ be a bounded vector field defined everywhere. For $t \in [0,T)$ fixed, we say that a map $Y_t : [0,t)\times\mathbb{T}^d \to \mathbb{R}^d$ is a backward regular Lagrangian flow with respect to the vector field b if:

1. *For $\mathscr{L}^d$-a.e. $x \in \mathbb{T}^d$ the map $s \mapsto Y_t(s,x)$ is an absolutely continuous integral solution of $\dot{\gamma}(s) = -b(t-s,\gamma(s))$ for $s \in [0,t)$, with initial condition $\gamma(0) = x$;*
2. *there exists a constant C independent of s,t such that*

$$Y_t(s,\cdot)_\#\mathscr{L}^d \le C\mathscr{L}^d \text{ as measures.}$$

Remark 2.2.4.
It is not difficult to see that $Y_t(s,X(t,x)) = X(t-s,x)$ just like in the classical case. Since we shall need a $L^1_t L^\infty_x$ bound on $\operatorname{div}_x(b)^-$ to have existence and uniqueness of the regular Lagrangian flow, in order to have also the existence and uniqueness of the backward regular Lagrangian flow we need to assume also that $\operatorname{div}_x(b)^+ \in L^1_t L^\infty_x$, i.e. that $\operatorname{div}_x(b) \in L^1_t L^\infty_x$.

We remark that Proposition 2.2.1 can be generalised to the case of regular Lagrangian flows, up to restricting the choice of the initial condition $\bar{\mu}$ accordingly.

Proposition 2.2.3.
Consider a vector field b which admits a regular Lagrangian flow $X : [0,T)\times\mathbb{T}^d \to \mathbb{R}^d$ and an initial density $\bar{\mu} = u_0 \in L^q(\mathbb{T}^d)$ for some $q \in [1,+\infty]$. Then

$$\mu_t = u(t,\cdot) := X(t,\cdot)_\# u_0$$

is a solution of (μ-CE) in the class $L^\infty(0,T;L^q(\mathbb{T}^d))$.

Proof. Identical to the proof of Proposition 1.2.1. □

We are now ready to state the following theorem, which relates the well-posedness of (μ-CE) in the class $L^\infty((0,T)\times\mathbb{T}^d)$ to the existence, uniqueness and stability of regular Lagrangian flows:

Theorem 2.2.2. ***Existence, uniqueness and stability of the regular Lagrangian flow, [57, Theorem 6.4.1 and Theorem 6.4.6]***
Let $b : [0,T)\times\mathbb{T}^d \to \mathbb{R}^d$ be a bounded vector field. The following hold:

- ***EXISTENCE**: If (μ-CE) admits a positive solution in $L^\infty((0,T)\times\mathbb{T}^d)$ with initial data $\bar\mu = \mathscr{L}^d$, then we have existence of a regular Lagrangian flow associated to b.*
- ***UNIQUENESS**: Assume that, if a positive solution $\mu_t \in L^\infty((0,T)\times\mathbb{T}^d)$ of (μ-CE) exists, it is unique. Then the regular Lagrangian flow associated to b, if it exists, is unique.*
- ***STABILITY**: Consider a sequence $\{b_k\}_{k\in\mathbb{N}}$ of vector fields such that*

$$\|b_k\|_{L^\infty((0,T)\times\mathbb{T}^d)} + \|\operatorname{div}_x(b_k)\|_{L^\infty((0,T)\times\mathbb{T}^d)} \le C$$

  for some $C>0$. Assume also that, for each b_k, (μ-CE) has the uniqueness property in $L^\infty((0,T)\times\mathbb{T}^d)$ and that the sequence $\{b_k\}_{k\in\mathbb{N}}$ converges in $L^1_{loc}((0,T)\times\mathbb{T}^d)$ to a vector field $b\in L^\infty((0,T)\times\mathbb{T}^d)$ with $\operatorname{div}_x(b)\in L^\infty((0,T)\times\mathbb{T}^d)$. Then the regular Lagrangian flows X_k associated to b_k converge strongly in $L^\infty(0,T;L^1_{loc}(\mathbb{T}^d))$ to the regular Lagrangian flow X associated to b [5].

Idea of the proof. We only give an idea on how to prove the uniqueness part, showing how the superposition principle can be employed. The key technical result is the following:

Proposition 2.2.4. ***[57, Proposition 6.4.3]***
Let b be a bounded vector field and assume that (μ-CE) has the uniqueness property in $L^\infty((0,T)\times\mathbb{T}^d)$. Consider a family of probabilities $\{\eta_x\}_{x\in\mathbb{T}^d}\subset\mathcal{P}(\Gamma_T)$ and assume that the measures η_x are supported on absolutely continuous integral solutions of (ODE) for $\mathscr{L}^d$ a.e. x. Then, if the superposition solution $\mu=\eta_x\otimes\bar\mu$ belongs to $L^\infty((0,T)\times\mathbb{T}^d)$, the measures η_x are Dirac deltas for $\mathscr{L}^d$-a.e. x.

We can argue by contradiction: consider two different regular Lagrangian flows X_1, X_2. Given an initial condition $\bar\mu\in L^\infty(\mathbb{T}^d)$ each of them induces bounded Lagrangian solutions μ_t^1,μ_t^2 of (μ-CE) which by the superposition principle are superposition solutions. By Proposition 2.2.4 the measures η_x^1,η_x^2 obtained by disintegrating μ_t^1,μ_t^2 are Dirac deltas for a.e. x. It is easy to prove that $\eta_x := \frac{1}{2}(\eta_x^1+\eta_x^2)$ induces the bounded solution $\mu_t := \eta_x\otimes\bar\mu$ of (μ-CE), and actually

$$\mu_t = \frac{1}{2}(\mu_t^1+\mu_t^2)\in L^\infty((0,T)\times\mathbb{T}^d), \tag{2.2.1}$$

which is a contradiction. Indeed, $\bar\mu$ and η_x satisfy the assumptions of Proposition 2.2.4, so for $\mathscr{L}^d$-a.e. x the measure η_x should be a Dirac mass, which is not the case. □

Remark 2.2.5. ***More general classes of vector fields***
If we consider vector fields with a particular structure, there are results also outside of the classes we previously considered. For instance, concerning the well-posedness of (ODE)*:*

- *In [44] it is proved existence and uniqueness of solutions if the vector field $b=(V,F(X))$ has an Hamiltonian structure and $F\in H^{3/4}(\mathbb{R}^d;\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d;\mathbb{R}^d)$.*
- *In [30] it is proved existence, uniqueness, quantitative stability and compactness when the gradient of the velocity field has the structure $D_x b = K * f$, with K a singular kernel and $f\in L^1(\mathbb{R}^d;\mathbb{R}^d)$, and they relate such a result to the well-posedness of Lagrangian solutions of the continuity equation.*

[5] This stability result seems very similar to Bressan's conjecture [32] However the crucial difference that makes the latter sensibly harder to prove is that in Bressan's conjecture we only have a L^1 uniform bound on the derivatives of b_k and no L^∞ control on $\operatorname{div}_x(b_k)$, but only on the compressibility of the flows.

- *In [130] quantitative stability estimates are given for vector fields of the form $b = K * g$ with $g \in L^1(\mathbb{R}^d;\mathbb{R}^d)$ and K a singular kernel.*

Remark 2.2.6. ***DiPerna-Lions definition of regular Lagrangian flow***
What we provided in Definition 2.2.4 is a definition due (at least, up to our knowledge) to Ambrosio [7]. The original definition of regular Lagrangian flow is due to DiPerna and Lions in [70] and it is slightly different. Most notably, they require a semigroup property for the flow (which actually follows from the two properties we used in Definition 2.2.4) and a lower bound on $X(t,\cdot)_\#\mathscr{L}^d$ which is completely absent from Ambrosio's definition. The idea is that "compressing" too much the measure is problematic (it might induce singular measures and, as we saw, staying inside the class where we have well-posedness for (μ-CE) is crucial), whereas "diluting" it is not. We refer the reader to [57, Section 6.5] for more details. We also remark that the strategy to prove well-posedness at the ODE level used in [70] is completely different and in a sense more standard, using approximation arguments.

Remark 2.2.7. ***General theory of $\mathscr{L}_b$ Lagrangian flows***
The assumptions in Theorem 2.2.2 are not sharp: for instance, only a control on the $L^1_t L^\infty_x$ norm of the negative part of the divergence of b is needed in the stability result. The interested reader can find sharper results in [10]. Moreover, in [10] a more general theory of so-called $\mathscr{L}_b$ Lagrangian flows is developed, of which regular Lagrangian flows are only an example. The basic idea is: if (μ-CE) is well posed in some class $\mathscr{L}_b \subset \mathcal{M}_+([0,T)\times\mathbb{T}^d)$, if we consider flows which solve (ODE), preserve the membership to $\mathscr{L}_b$ (i.e. if $\bar\mu \in \mathscr{L}_b \implies X(t,\cdot)_\#\bar\mu \in \mathscr{L}_b$) and the class $\mathscr{L}_b$ is convex (so that we can argue as in (2.2.1)), then uniqueness of such flows can be proved.

Remark 2.2.8. ***Almost everywhere uniqueness of integral curves for regular Lagrangian flows***
It is quite natural to wonder if, given the unique regular Lagrangian flow X of a vector field b, it holds that for almost every x there is a unique integral curve of b satisfying $\gamma'(t) = b(t,\gamma(t))$ with $\gamma(0) = x$. We remark that the uniqueness of the regular Lagrangian flow that we proved does not imply almost everywhere uniqueness of integral curves. Actually, such uniqueness does not hold in general. In [36] the authors showed that there exists a vector field $b \in C([0,1];W^{1,r}(\mathbb{T}^d;\mathbb{R}^d))$ with $r < d$ which admits a unique regular Lagrangian flow X such that, if $x \in A$ with A a set of positive measure, there are at least two integral curves of b starting from x. See also [137] for a slight generalization. If instead $r > d$ such almost everywhere uniqueness holds, as shown in [43, Corollary 5.2]. We also cite [102] where the author constructs a continuous vector field (both in time and space) for which uniqueness of integral curves does not hold in a set of positive Lebesgue measure, answering a question posed by Alberti in [4]. Such a construction will be the basis of [37], where the authors develop a new method to build non-unique solutions of (TE).

Remark 2.2.9. ***Is the positivity assumption really necessary?***
It is natural to wonder if the positivity assumption in the superposition principle is really necessary or if it is just a shortcoming of the theory. As it turns out, in general a signed superposition principle does not hold simply by observing that there exist signed solutions of (μ-CE) even for vector fields b which do not have solutions for (ODE) *for any $x \in \mathbb{R}^d$ [87]. However, assuming some regularity on the vector field b, a signed superposition principle does hold [9]. One may also wonder if the implication $1. \implies 2.$ in Proposition 2.2.2 really needs the non-negativity assumption. In general, the implication is false [28].*

2.2.2 Applications to positive solutions of the continuity equation

The superposition principle (Theorem 2.2.1) can be used not only as a tool to prove well-posedness of (ODE) based on the well-posedness of (μ-CE). It can also be used to go beyond the DiPerna-Lions framework, even though for positive solutions only. The main result in this setting is the following:

Theorem 2.2.3. ***Uniqueness of positive solutions, [36]***
Let $d \geq 2$, $p \in [1,+\infty]$ and $r \in [1,+\infty]$ be such that

$$\frac{1}{p} + \frac{1}{r} < 1 + \frac{1}{d-1}\frac{r-1}{r}.$$

Let $b \in L^1(0,T;W^{1,r}(\mathbb{T}^d;\mathbb{R}^d))$ be a vector field satisfying $\operatorname{div}_x(b) \in L^\infty((0,T)\times\mathbb{T}^d)$. Then, for $p > 1$, (μ-CE) admits a unique solution among all non-negative, weakly continuous in time densities ρ in the class

$L^\infty(0,T;L^p(\mathbb{T}^d;\mathbb{R}^+_0))$ with $\rho(0,\cdot)=\rho_0$. When $p=1$ (i.e.$r>d$) uniqueness holds in the class of non-negative weakly-$*$ continuous densities $\rho\in L^\infty(0,T;\mathcal{M}^+(\mathbb{T}^d))$ with $\rho(0,\cdot)=\rho_0\mathscr{L}^d$. In particular every such ρ is Lagrangian, i.e.

$$\rho(t,\cdot)\mathscr{L}^d = X(t,\cdot)_\#(\rho_0\mathscr{L}^d),$$

where X denotes the unique regular Lagrangian flow of b.

2.3 Non-uniqueness phenomena

We now briefly turn our attention to counterexamples, showing how the assumptions in the previous sections are basically sharp. We will not give the full details of the counterexamples, referring the reader to the relevant literature for more insight. As we have already done repeatedly, in this section we will also move back and forth between the Lagrangian and the Eulerian points of views, showing non-uniqueness phenomena for both (ODE) and (CE),(TE).

2.3.1 Beck/DiPerna-Lions and Aizenman/Depauw counterexamples

We start this discussion from a Lagrangian point of view, i.e. looking for non-uniqueness of solutions of (ODE). Notice that we are not asking for non-uniqueness of a Cauchy problem, otherwise Example 1.1.1 would have been sufficient. We are looking for non-uniqueness of flows[6] which satisfy (ODE). Historically the first counterexamples are given in [17, 3]. In [3] in particular the author constructs uniformly bounded divergence-free vector fields b with any of the following properties:

- There is no measure-preserving flow generated by b;
- the vector field b generates an uncountable number of flows;
- along integral lines, fibers are contracted to points and vice-versa points burst into fibers.

Later, two counterexamples are provided in [70]. Given $p\in[1,+\infty)$, the first counterexample consists in an autonomous vector field $b\in[W^{1,p}_{loc}(\mathbb{R}^2)\cap L^\infty(\mathbb{R}^2)]^2$ with unbounded divergence. Based on the construction in [17] they define an infinite number of flows X_μ parametrized by positive measures in a Cantor set. It is worth noting that each one of these flows concentrate absolutely continuous measures in singular ones, except for the flow corresponding to the measure $\mu=0$. This means that this is not a counterexample for the uniqueness of regular Lagrangian flows. The second counterexample concerns an autonomous divergence-free vector field in $\mathbb{R}^2$ with fractional order derivatives (see [1, 69] for some general theory on fractional Sobolev spaces). Namely, $b\in[W^{s,1}_{loc}(\mathbb{R}^2)]^2$ for every $s\in[0,1)$ and $b\in[L^\infty(\mathbb{R}^2)\cap L^p(\mathbb{R}^2)]^2$ for every $p\in[1,2)$. In this case, the authors construct two different measure-preserving regular Lagrangian flows relative to this vector field. As a byproduct, this is also a counterexample to the uniqueness of solutions of (CE) when $\bar u\in L^\infty(\mathbb{T}^d)$: the push-forward with respect to the two different regular Lagrangian flows gives two different bounded solutions of (CE). The next counterexample, up to our knowledge, was provided (independently from each other) in [66] and [49], both building on and simplifying the idea of [3]. In particular the construction in [66] is very simple and short (only four pages long). Contrary to the counterexamples in [70] and the original work [3], the focus in [66] is on the Eulerian point of view. Namely, the author exhibits a non-zero bounded solution of (CE) with a zero initial datum. What is truly remarkable is how close the vector field b considered in [66] is to satisfying the assumptions in Theorem 2.1.5: it is bounded, divergence-free and, for every $\varepsilon>0$, $b\in L^1(\varepsilon,T;BV_{loc}(\mathbb{R}^2;\mathbb{R}^2))$, but $b\notin L^1(0,T;BV_{loc}(\mathbb{R}^2;\mathbb{R}^2))$ (otherwise uniqueness for (CE) would hold by Theorem 2.1.5). For the definition of the vector fields and a precise analysis of the counterexamples in [70, 66], other than the original articles we refer the reader to [57].

[6] In Definition 1.1.1 we never assumed the flow to have a semigroup property, i.e. that $X(s,X(t,x))=X(s+t,x)$, because it followed by our assumptions on b. In general, the semigroup property is part of the definition of a flow, as in [17, 3]. It follows that, in general, defining $X(t,x)$ as the value at time t of an integral curve starting at x does not necessarily produce a flow, since such curves may not be unique and may yield a map X lacking the semigroup property.

2.3.2 Modena-Sattig-Szekelyhidi: convex integration

The convex integration method was initially introduced in [64, 63] in the context of fluid dynamics to provide counterexamples for Euler and Navier-Stokes equations [59, 38, 39], most notably resulting in proving a conjecture by Onsanger [132] in [95] and in proving non-uniqueness of weak solutions for 3D Navier-Stokes with finite kinetic energy [40]. For more details on the convex integration scheme in fluid dynamics, we refer the reader to [41]. In the seminal works [126, 127, 125] the method was adapted to provide counterexamples to uniqueness of bounded solutions for (CE), ultimately providing the following result:

Theorem 2.3.1. ***[125]***
Let $q \in [1, +\infty)$, $p \in [1, +\infty)$ and assume that

$$\frac{1}{p} + \frac{1}{q} > 1 + \frac{1}{d}.$$

Then there are infinitely many incompressible vector fields

$$b \in [C_t L^{q'}(\mathbb{T}^d) \cap C_t W^{1,p}(\mathbb{T}^d)]^d$$

for which uniqueness of (CE) *fails in the class of densities $\rho \in C_t L^q(\mathbb{T}^d)$. Moreover, if $q = 1$, it holds $b \in C([0,T] \times \mathbb{T}^d; \mathbb{R}^d)$.*

Before giving the basic idea of the convex integration scheme applied to the transport equation, let us remark that the convex integration method is a purely Eulerian technique. Namely, there is no flow involved and we are actually in a setting where the regular Lagrangian flow of b is unique [7]. The starting point is to notice the following fact: if ρ is a solution of (CE) and the vector field is incompressible, it means it satisfies

$$\int_0^T \int_{\mathbb{T}^d} \rho(\partial_t \varphi + b \cdot \nabla_x \varphi) dx dt + \int_{\mathbb{T}^d} \varphi(0,x)\rho_0(x) dx = 0 \text{ for every } \varphi \in C_c^\infty([0,T) \times \mathbb{T}^d).$$

We have seen that if $b \in [L^1(0,T; W^{1,p}(\mathbb{T}^d)) \cap L^\infty((0,T) \times \mathbb{T}^d)]^d$, solutions ρ of (CE) exist and are unique in the class $L^\infty(0,T; L^q(\mathbb{T}^d))$ with $1/p + 1/q \leq 1$. The above equation however makes sense even for $q' < q$, because by Sobolev embedding $b \in L^{p^*}(\mathbb{T}^d; \mathbb{R}^d)$, with $p^* = dp/(d-p)$. The $W^{1,p}$ regularity in particular is needed only for proving the renormalization property (Lemma 2.1.1), while for existence it is not needed at all. It is then natural to wonder whether the class where we have uniqueness can be extended from $L^\infty(0,T; L^q(\mathbb{T}^d))$ to some $L^\infty(0,T; L^{q'}(\mathbb{T}^d))$ with $q' < q$. There is a competition: on the one hand the Lagrangian setting is very well behaved since there is a unique regular Lagrangian flow and thus there is only one Lagrangian solution of (CE), on the other hand the lower regularity of the density might break the DiPerna-Lions renormalization scheme and there might exist weak solutions which are not renormalized. As we saw in Theorem 2.2.3, if we restrict to positive solutions we have uniqueness even outside the DiPerna-Lions setting, but in general Theorem 2.3.1 shows that we have not.

We now (very briefly) explain the idea behind the convex integration, which is an iterative scheme. We consider a couple (ρ_n, b_n) solving:

$$\partial_t \rho_n + \operatorname{div}_x(b_n \rho_n) = -\operatorname{div}_x(R_n),$$

where $\operatorname{div}_x(R_n)$ can be thought as an error term. The goal is to define a new pair (ρ_{n+1}, b_{n+1}) satisfying

$$\partial_t \rho_{n+1} + \operatorname{div}_x(b_{n+1}\rho_{n+1}) = -\operatorname{div}_x(R_{n+1})$$

such that the new error term is smaller, and in particular we want that $\operatorname{div}_x(R_n) \rightharpoonup 0$ as $n \to +\infty$ so that in the limit (ρ_n, b_n) converge to a couple (ρ, b) solving (CE). The way to define ρ_{n+1} and b_{n+1} starting from ρ_n, b_n is by adding a perturbation term, namely:

$$\rho_{n+1} := \rho_n + \varepsilon_n \text{ and } b_{n+1} := b_n + \delta_n.$$

[7]In the previous counterexamples the Lagrangian point of view was used to show the non-uniqueness also at the PDE level, and this was achieved by considering vector fields for which (ODE) admits two different regular Lagrangian flow, with which one can define (simply by push-forward of the initial datum) two different solutions of (CE). That is: non-uniqueness of (ODE) $\Longrightarrow$ non-uniqueness of (CE).

For the moment let us not specify ε_n and δ_n. If b_{n+1} and ρ_{n+1} are defined like above, then they satisfy:

$$\partial_t \rho_{n+1} + \operatorname{div}_x(\rho_{n+1} b_{n+1}) = \operatorname{div}_x(\varepsilon_n \delta_n - R_n) + \text{ other error terms.}$$

The crucial point is to choose ε_n and δ_n in a way to make the error term $\operatorname{div}_x(\varepsilon_n \delta_n - R_n)$ smaller and smaller and a way this can be achieved is by considering ad hoc functions (which we will not specify) such that

$$\operatorname{div}_x(\varepsilon_n \delta_n - R - n) = f g_n,$$

where $g(x)$ satisfies $\int_{\mathbb{T}^d} g(x)dx = 0$, $g_n(x) := g(\lambda_n x)$ and $\lambda_n \to +\infty$. Then, it is easy to see that $f g_n \rightharpoonup f \fint_{\mathbb{T}^d} g(x)dx = 0$. The key point is to add functions with higher and higher frequencies to reduce the error term. Of course, the argument is much more sophisticated than this, and we refer the reader to [125] for the details, but it is relevant (especially with respect to the next subsection) to notice that at each step the perturbations are added uniformly in space, i.e. the functions ε_n and ρ_n are approximately $1/\lambda_n$ periodic. In the next subsection we will see how the range for non-uniqueness of Theorem 2.3.1 can be improved by considering both a different type of flow of b (i.e. not a regular Lagrangian flow) and using non-uniform perturbations.

2.3.3 Bruè-Colombo-Kumar: a (non-regular) Lagrangian flow approach

We conclude the topic of non-uniqueness discussing the very recent work [37], returning once again to a Lagrangian point of view. As we already said, if a vector field b has Sobolev space regularity and some bounds on the divergence, then solutions of (CE) are unique in $L^\infty((0,T)\times\mathbb{T}^d)$ and the regular Lagrangian flow of b is also unique. However, if we consider the class $L^\infty(0,T;L^q(\mathbb{T}^d))$ with $q \neq \infty$ there might be flows satisfying (ODE) for almost every x such that they are not the regular Lagrangian flow of b and yet, given an initial datum $\bar\rho \in L^q(\mathbb{T}^d)$, the Lagrangian solution $X(t,\cdot)_\# \bar\rho$ obtained by push-forward of the initial datum is uniformly bounded in $L^\infty(0,T;L^q(\mathbb{T}^d))$. Namely, there might be flows which are sort of intermediate between the regular Lagrangian flow (which preserves boundedness in L^q if the initial datum is in L^q for every $q \in [1,+\infty]$) and flows like the one presented in the first counterexample of [70], where absolutely continuous initial data are concentrated in singular measures. The goal is thus to construct a flow X which solves (ODE) that is not the regular Lagrangian flow of b, but which still keeps the L^q norm of the solution $X(t,\cdot)_\#\bar\rho$ uniformly bounded. The strategy of [37] is the following: using a construction based on the one proposed in [102] they construct a (sort of) self-similar solution that belongs to $L^\infty_t L^q_x$ (for a range of p,q that will be specified shortly) and which starts from the uniform distribution on a cube but which differs from the one obtained via the regular Lagrangian flow (which is simply the constant one, since the vector field they construct is divergence-free and supported on the aforementioned cube). We give the statement of the main result of [37].

Theorem 2.3.2. ***[37]***
Let $d \geq 2$, $p \in [1,+\infty)$, $q \in [1,+\infty)$ be such that

$$\frac{1}{p} + \frac{1}{q} > 1 + \frac{1}{d-1}\frac{p-1}{p}.$$

Then there exists a compactly supported divergence-free vector field $b \in C([0,1]; W^{1,p} \cap L^{\frac{q}{q-1}}(\mathbb{T}^d;\mathbb{R}^d))$ such that (CE) *admits two distinct non-negative solutions in the class $C([0,1];L^q(\mathbb{T}^d))$.*

We now briefly explain the idea behind this result and for simplicity we will restrict to the planar case and, for the moment, $q = 1$. We will use Figure 2.1, taken from [37], as a visual aid. Consider as initial datum the four squares in (a), where the bigger square they are contained in is $Q = [-1/2, 1/2]^2$. At each step we want to break each one of the four squares into four smaller squares, contract them to increase the density (but the total mass, i.e. the L^1 norm will always remain constant) and separate them in order to have smaller copies of the original problem. These are respectively the steps (b) and (c) in Figure 2.1. Now we have four copies of the original problem and we iterate this procedure at higher and higher spatial and time frequencies (namely, performing this splitting for smaller and smaller squares requires less and less time). By iterating this procedure we can make the initial density concentrate more and more on a Cantor-like set where the density explodes to $+\infty$, but the measure of the support goes to 0. If we choose carefully the compression rate of the squares we can prove that in the limit $t \to 1$ the density converges to the uniform distribution in Q. It can also be proved that the vector field v which provides this "cut, compress and split" procedure is divergence-free, supported in Q and $v \in C(0,1; W^{1,p} \cap L^{\frac{q}{q-1}}(\mathbb{T}^2;\mathbb{R}^2))$, so there exists a unique regular Lagrangian flow relative to it which is measure preserving. Consider now $b(t,x) := v(1-t,x)$ the time-reversal of the vector field. If we consider $\bar{\rho} = \mathbb{1}_Q$ as initial datum, then $\rho \equiv \bar{\rho}$ is a solution of (CE), since the regular Lagrangian flow of b is measure preserving and supported in Q. At the same time, the density we constructed iteratively is also a solution, which by construction belongs to $C(0,1; L^1(\mathbb{T}^2))$, so we have proved non-uniqueness. If $q > 1$, there is an additional problem. We cannot perform the same steps we used in the $q = 1$ case because we would have that, for every $q > 1$, $\sup_{t\in[0,1]} \|\rho(t,\cdot)\|_{L^q} = +\infty$. The problem is that the uniform (in space) splitting that worked in the L^1 case produces a blow-up of the L^q norms for every $q > 1$. A way to solve the problem, and a fundamental difference with respect to the convex integration method (other than one being fundamentally Lagrangian and the other one fundamentally Eulerian), is to perform asynchronous steps in the splitting procedure, not focusing on all the "big squares" at the same time, but considering only one at a time, choosing smaller and smaller ones. We refer to Figure 2.2 for a visual explanation. A good analogy which helps understanding the concept is that of a tree where the first four squares in (a) of Figure 2.1 are the nodes of first layer, the second layer is composed of the sixteen smaller squares in (c) and so on. The goal is for us to visit each node of the tree, where with each visit we split the square into smaller and smaller ones which in the limit gives us the constant density $\bar{\rho}$. While in the L^1 case we could perform a breadth-first search, in the L^q case a depth-first visit is needed. We can see it in Figure 2.2: the splitting always occurs on the bottom left smaller and smaller square. Using faster and faster splittings the smaller the (single!) square gets, we can recover the convergence to $\bar{\rho}$, without blowing up the L^q norm, even with this asynchronous procedure. We refer the reader to [37] for the rigorous proof of the result and all the details.

Remark 2.3.1. ***The stochastic counterpart***
Before concluding, it is important to remark that problems involving the well-posedness of (ODE) *have been studied also in the setting of stochastic differential equations (SDEs). As a general fact, it is well-known that noise can prevent non-uniqueness phenomena, for instance we mention the (by now classical) result of Veretennikov [156]: the SDE $dX_t = b_t(X_t)dt + dW_t$ has strong existence and pathwise uniqueness if b is merely bounded* [8]. *Later, in [101] existence and uniqueness of strong solutions of the SDE was obtained by requiring a Ladyzhenskaya-Prodi-Serrin integrability condition, namely:*

$$b \in L^q_t L^p_x \text{ such that } \frac{d}{p} + \frac{2}{q} < 1,\ p \geq 2,\ q > 2.$$

Finally, we cite [45] where they focus on SDEs with a Sobolev drift and a multiplicative noise, proving well-posedness in that setting. There are also works that relate the well-posedness of the SDE with the well posedness of associated PDEs (such as the Fokker-Planck equation, see Chapter 6 for more details). One of the first works in this direction, up to our knowledge, is [77], where the author proves a stochastic counterpart of Ambrosio's theory on regular Lagrangian flows, however it still assumes BV regularity for the vector field and a $L^1_t L^\infty_x$ condition on the divergence. A more refined version of the superposition principle in the stochastic setting was later proved in [154]. A breakthrough on another PDE associated to the SDE (namely a stochastic transport equation, i.e. (TE) *with the addition of a multiplicative random perturbation term) was obtained in [78] where the authors managed to prove well posedness under the assumption that the velocity field is bounded, it has uniformly bounded Hölder regularity (in space) and the divergence is an L^p function in $[0,T] \times \mathbb{R}^d$ for a suitable p.*

[8] There are various different notions of existence and uniqueness in the context of SDEs. In particular, we refer the reader to [15] for the definitions of weak/strong solutions and to [146] for the different types of uniqueness.

For the reader's convenience, we conclude this chapter with tables that summarise the uniqueness results, both at the ODE and at the PDE level, which are currently known.

	$1/p + 1/q \leq 1$	$1 < 1/p + 1/q < 1 + \frac{p-1}{(d-1)p}$	$1/p + 1/q > 1 + \frac{p-1}{(d-1)p}$
Signed solutions	Uniqueness [70]	Open problem	Non uniqueness [37]
Non-negative solutions	Uniqueness [70]	Uniqueness [36]	Non uniqueness [37]

Table 2.1: Range of uniqueness in the class $L^\infty(0,T;L^q(\mathbb{T}^d))$ for solutions of (CE) and (TE), assuming $b \in [L^1(0,T;W^{1,p} \cap L^\infty(\mathbb{T}^d))]^d$ with $\mathrm{div}_x(b) \in L^\infty((0,T) \times \mathbb{T}^d)$.

	$p < d$	$p > d$
Almost everywhere uniqueness of trajectories	Non uniqueness [36, 137, 102]	Uniqueness [43]

Table 2.2: Range of uniqueness of trajectories for almost every x if b has spatial $W^{1,p}$ regularity.

Remark 2.3.2. ***Some caveats***
We recall that, if $q = 1$, the uniqueness for non-negative solutions holds in the larger class of measure valued functions $L^\infty(0,T;\mathcal{M}(\mathbb{T}^d))$ [36]. If $d = 1$ or $d = 2$ the situation is described in detail in Remark 2.1.7. Moreover, up to our knowledge, the critical cases $\frac{1}{p} + \frac{1}{q} = 1 + \frac{p-1}{(d-1)p}$ for non-negative solutions of (CE), (TE) *and the case $p = d$ for* (ODE) *are completely open.*

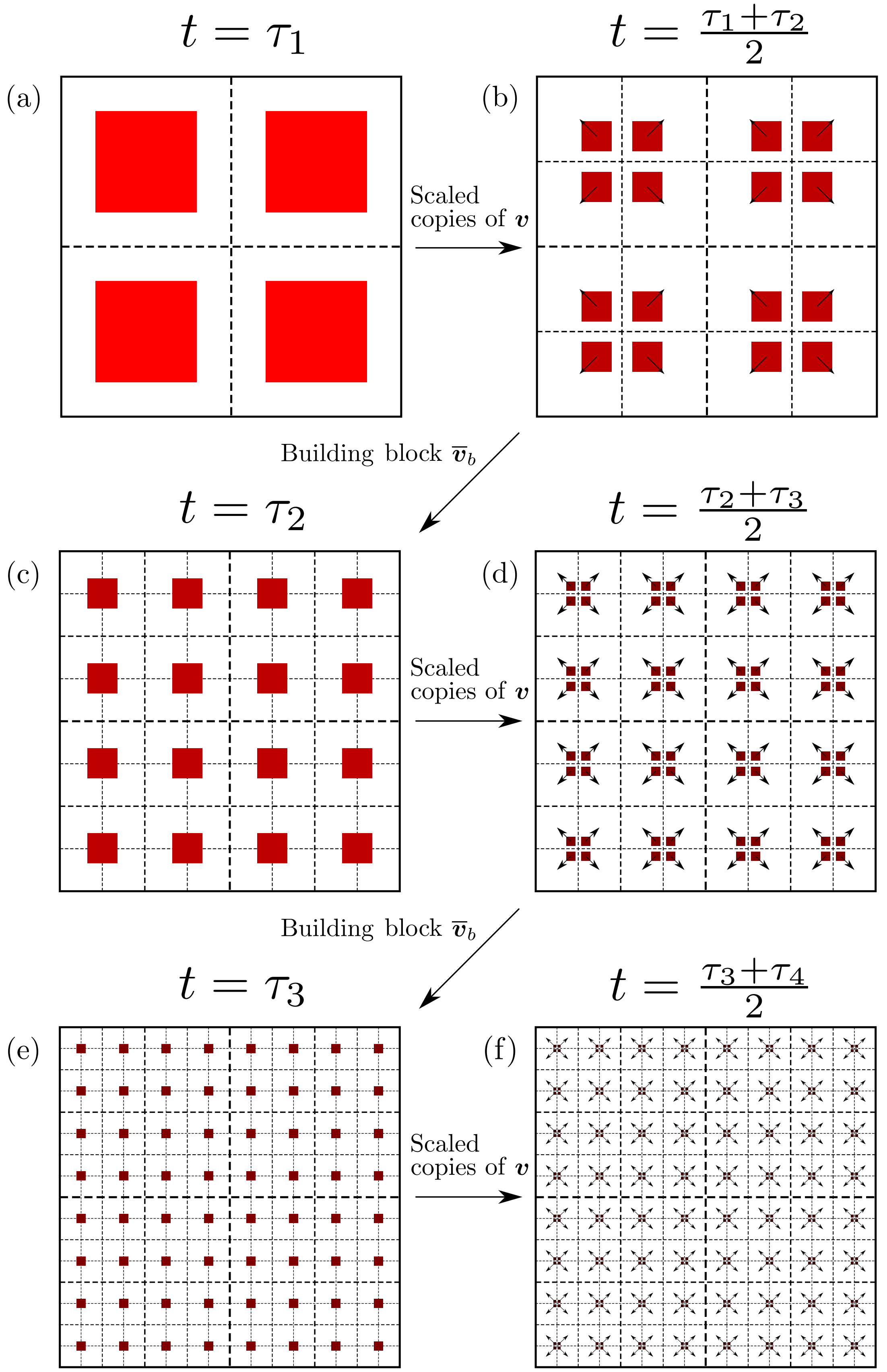


Figure 2.1: Construction in [37] for $q = 1$. A darker colour denotes a higher density. Figure kindly provided by prof. Maria Colombo.

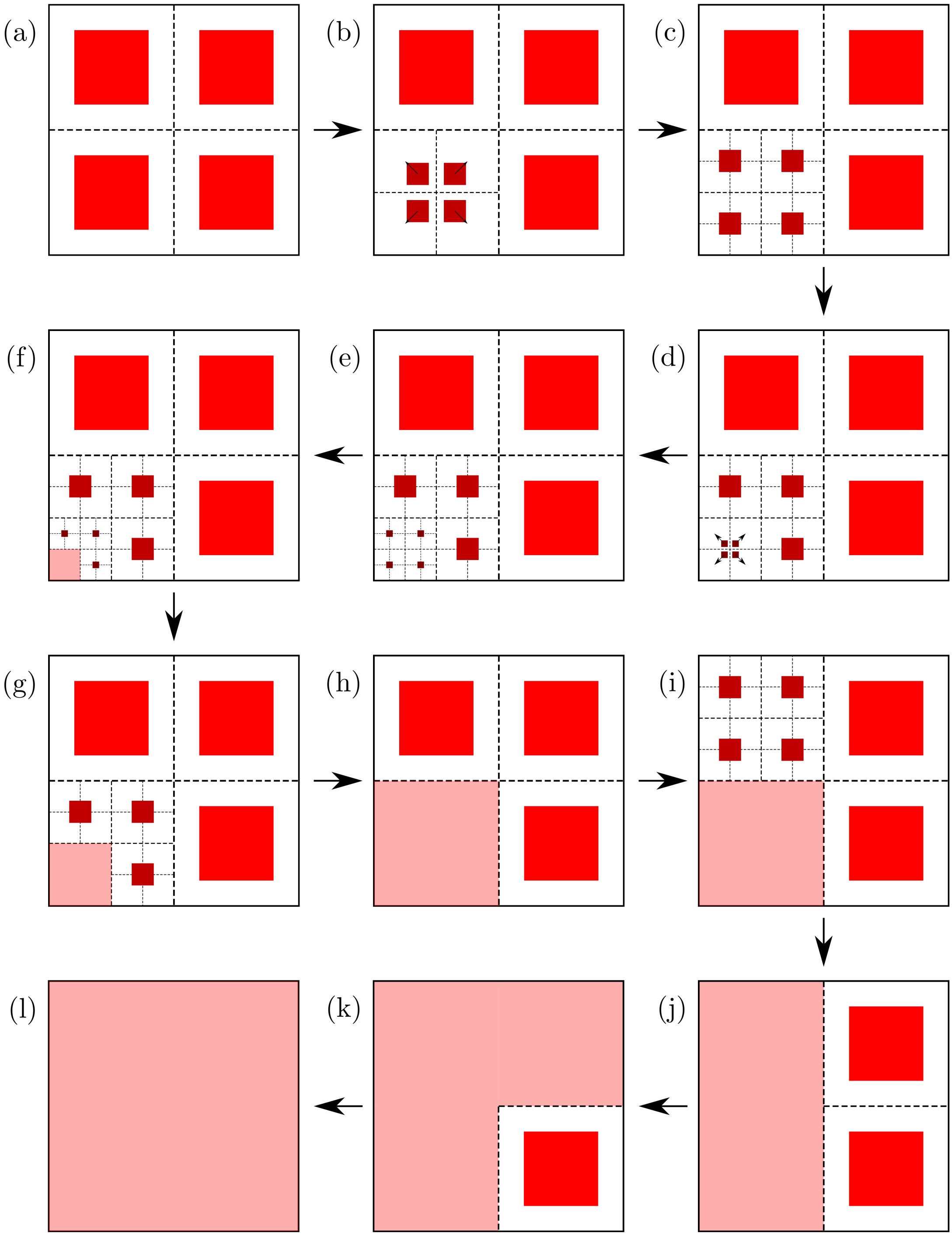


Figure 2.2: Construction in [37] for $q > 1$. A darker colour denotes a higher density. Figure kindly provided by prof. Maria Colombo.

Chapter 3

A more geometric (measure theory) approach

In this chapter we want to prove the well-posedness of the continuity and of the transport equations without relying on the theory of renormalized solutions, but instead via a more geometric approach reminiscent of the classical method of characteristics, using the language of currents borrowed from geometric measure theory. The idea is to view the continuity equation in an equivalent way as a condition on the boundary of a 1-current with a particular structure. Then the uniqueness follows by an approximation argument using smooth vector fields and by a very simple (almost trivial) result: if the vector field is zero, then the mass of the current is controlled by the mass of the initial datum. That is: if the velocity field is zero, the unique solution is the constant one. The main technical point will however still be a strong L^1 estimate on the commutator, so this different point of view does not seem to yield any technical advantage over the renormalization approach, at least for the moment. We point out that a different method for proving the uniqueness of the continuity equation without using the renormalization technique has also been given in [144], in which the author essentially relied on Lagrangian stability estimates provided in [58] and on results from Optimal Transport theory. We also notice that recently a link between currents and the continuity equation has also been proposed in [27], where the authors prove the well-posedness of a generalisation to k-currents of the continuity and of the transport equations in $\mathbb{R}^d$ (which are obtained in the particular cases of $k = 0$ and $k = d$, respectively). For these reasons, we believe that a more geometric point of view on the subject may be of interest for future developments. We recall for convenience that the continuity equation we will consider in the following is

$$\begin{cases} \partial_t u + \operatorname{div}_x(bu) = 0 & \text{in } \mathbb{T}^d \text{ with } t \in (0,1) \\ u(0,\cdot) = u_0 \end{cases}.$$

Throughout this chapter we will assume without loss of generality that the final time is 1. This is clearly not restrictive and it will avoid ambiguity since the letter T we used in previous chapters will be reserved for another object, namely a normal 1-current representing the solution of the continuity equation. We restrict to the case of Sobolev bounded velocity fields defined in $\mathbb{T}^d$ for simplicity, but as noted in Remark 3.2.5 the same argument can be adjusted to work in the BV case, too. As in other chapters, we work in the torus $\mathbb{T}^d$ instead of $\mathbb{R}^d$ to avoid localisation arguments and we use Einstein's summation convention, i.e. we implicitly sum over repeated indices.

3.1 Preliminary definitions

In order to keep this chapter as self-contained as possible, we give here the necessary definitions. Since we mainly work with 1-currents, we will only define them, but the interested reader may find more general definitions in [75, 148]. Definitions are given in $\mathbb{R}^d$, but they carry over to $\mathbb{T}^d$ without any modification. In the following we will always assume $U, V \subset \mathbb{R}^d$ to be open sets.

Definition 3.1.1. ***1-currents***
A 1-current on U is a continuous linear functional on the space $\mathcal{D}^1(U)$ of 1-differential forms on U. That is, 1-currents are elements of the dual of $\mathcal{D}^1(U)$, and this set is denoted as $\mathcal{D}_1(U)$.

Definition 3.1.2. ***Pull-back of a differential form and push-forward of a current***
Let $\omega(x) = \sum_{1\le i\le d} \xi_i(x)dx_i$ be a 1- differential form in $\mathcal{D}^1(V)$, with $\xi_i \in C_c^\infty(V)$ for every i, and let $\phi : U \mapsto V$ be a smooth map with compact support. We denote the pull-back of ω via ϕ by $\phi^\#\omega \in \mathcal{D}^1(U)$, which is defined as

$$\phi^\#\omega(x) := \sum_{1\le i\le d} \xi_i(\phi(x))\, d\phi_i,$$

with $d : C_c^\infty(V) = \mathcal{D}^0(V) \to \mathcal{D}^1(V)$ the differential operator. Similarly, given a 1- current $T \in \mathcal{D}_1(U)$, we denote the push-forward of T via ϕ by $\phi_\# T \in \mathcal{D}_1(V)$, and it is defined as

$$\phi_\# T(\omega) := T(\phi^\#\omega) \quad \text{for all } \omega \in \mathcal{D}^1(V).$$

Definition 3.1.3. ***Boundary of a 1-current***
Let $T \in \mathcal{D}_1(U)$. We define the boundary of T as a distribution on U (i.e. an element of the dual space of $\mathcal{D}^0(U) = C_c^\infty(U)$) as

$$\partial T(\xi) := T(d\xi) \quad \text{for all } \xi \in C_c^\infty(U).$$

Remark 3.1.1. ***Properties of push-forward and boundary & of pull-back and differential***
It can be verified that the pull-back and the differential commute, i.e. if we consider a differential form ω and a pull-back via a smooth map ϕ, then

$$\phi^\#(d\omega) = d(\phi^\#\omega).$$

Similarly, the boundary operator commutes with the push-forward for currents, i.e. for every current T and for every smooth map ϕ it holds that

$$\phi_\#(\partial T) = \partial(\phi_\# T).$$

Definition 3.1.4. ***Mass of a 1-current***
Let $T \in \mathcal{D}_1(U)$. We define the mass of T as

$$\mathbb{M}(T) = \sup_{\substack{\omega\in\mathcal{D}^1(U) \text{ s.t.} \\ |\xi_i|\le 1 \text{ for every } i}} T(\omega), \qquad \text{with } \omega(x) = \sum_{1\le i\le d} \xi_i(x)dx_i.$$

If S is a distribution in U (i.e. $S \in \mathcal{D}_0(U)$), we define its mass as

$$\mathbb{M}(S) := \sup_{\substack{\xi\in C_c^\infty(U) \text{ s.t.} \\ |\xi|\le 1}} S(\xi).$$

Definition 3.1.5. ***Normal 1-currents***
We say that a 1-current $T \in \mathcal{D}_1(U)$ is normal if

$$\mathbb{M}(T) + \mathbb{M}(\partial T) < +\infty.$$

An example of a 1-current T is

$$T = \sum_{1\le i\le d} f_i e_i,$$

with $f_i \in L^1_{loc}(U)$ for every i. In this case, the action of T on a 1-form $\omega(x) = \sum_{1\le i\le d} \xi_i(x)dx_i \in \mathcal{D}^1(U)$ is defined as

$$T(\omega) := \int_U \left(\sum_{1\le i\le d} f_i(x)\xi_i(x) \right) dx.$$

Definition 3.1.6. ***Flat norm***
Let $T \in \mathcal{D}_1(U)$. We define the flat norm of T as

$$\mathbb{F}(T) := \inf\{\mathbb{M}(S) + \mathbb{M}(L) \text{ s.t. } T = \partial S + L\}.$$

In particular, we point out that S is a 2-current (see [75, 148]) and that $\mathbb{F}(T) = 0 \implies T = 0$.

3.2 Proof of uniqueness

Let $1 \le p, q \le +\infty$ such that $1/p + 1/q \le 1$. In this chapter we will always work with the following hypotheses on the vector field b and on the initial datum u_0:

$$\begin{cases} b \in [L^1(0,1; W^{1,p}(\mathbb{T}^d)) \cap L^\infty((0,1) \times \mathbb{T}^d)]^d \\ \operatorname{div}_x(b)^- \in L^1(0,1; L^\infty(\mathbb{T}^d)) \\ u_0 \in L^q(\mathbb{T}^d) \end{cases} . \tag{Hyp}$$

We will prove uniqueness of solutions in the class

$$u \in L^\infty(0,1; L^q(\mathbb{T}^d)).$$

The starting point of our approach is to notice that having a weak solution u of the continuity equation can be written in an equivalent formulation in the language of currents.

Definition 3.2.1. ***Solution in the sense of currents***
A function $u \in L^\infty(0,1; L^q(\mathbb{T}^d))$ is said to be a solution in the sense of currents of the continuity equation if the associated normal 1-current T in $(-\infty, 1) \times \mathbb{T}^d$ and supported in $[0,1) \times \mathbb{T}^d$, defined as $T = ue_t + ub^i e_i$, satisfies

$$\partial T = -u_0 \mathcal{H}^d \llcorner (\{0\} \times \mathbb{T}^d), \tag{C-CE}$$

where the boundary is considered with respect to $(-\infty, 1) \times \mathbb{T}^d$; e_t is the unitary vector in the time direction; $e_1, e_2, \dots, e_d$ are unitary vectors in the x_i directions and the b^i's are the components of b.

Remark 3.2.1.
Notice that we will always consider currents in $(-\infty, 1) \times \mathbb{T}^d$ and supported in $[0,1) \times \mathbb{T}^d$. The reason of this mismatch is that the definition of the boundary of a current is better given on open sets.

Remark 3.2.2.
Let us check that Definition 3.2.1 is equivalent to the notion of weak solution of the continuity equation given in (2.1.1)*, i.e. u satisfies the continuity equation if and only if it satisfies* (C-CE)*. Assume that u satisfies* (2.1.1) *and consider its associated current $T = ue_t + ub^i e_i$. We have:*

$$\langle \partial T, \xi \rangle = \langle T, d\xi \rangle = \int_0^1 \int_{\mathbb{T}^d} u[\partial_t \xi + b \cdot \nabla_x \xi] dx dt \qquad \forall \xi \in C_c^\infty([0,1) \times \mathbb{T}^d).$$

Integrating by parts the first addendum:

$$\begin{aligned} \int_0^1 \int_{\mathbb{T}^d} u \partial_t \xi dx dt &= \int_{\mathbb{T}^d} \left([u\xi]_0^1 - \int_0^1 \xi \partial_t u dt \right) dx \\ &= -\int_{\mathbb{T}^d} u_0 \xi dx - \int_0^1 \int_{\mathbb{T}^d} \xi \partial_t u dx dt. \end{aligned}$$

As for the second addendum, we use that

$$\operatorname{div}_x(ub\xi) = \operatorname{div}_x(ub)\xi + ub \cdot \nabla_x \xi \text{ in the sense of distributions,}$$

so

$$\int_0^1 \int_{\mathbb{T}^d} ub \cdot \nabla_x \xi dx dt = -\int_0^1 \int_{\mathbb{T}^d} \xi \operatorname{div}_x(bu) dx dt.$$

Using that by assumption u satisfies the continuity equation, we easily see that

$$\langle \partial T, \xi \rangle = -\int_{\mathbb{T}^d} u_0 \xi dx.$$

If we consider $T = ue_t + ub^i e_i$ satisfying (C-CE)*, by the same argument u also satisfies* (2.1.1)*, i.e. it is a solution of the continuity equation.*

We will consider a 1-current $T = ue_t + ub^i e_i$ satisfying (C-CE) with null boundary in $(-\infty, 1) \times \mathbb{T}^d$ (i.e. $u_0 = 0$), with b satisfying (Hyp), and we shall prove that $T = 0$. This is equivalent to proving uniqueness of solutions of the continuity equation by linearity of the equation and Remark 3.2.2. The main result of the chapter, which implies uniqueness, is the following:

Theorem 3.2.1. ***Mass estimate***
Let b satisfy (Hyp) *and let $T = ue_t + ub^i e_i$ be the 1-current associated to a density $u \in L^\infty(0,1; L^q(\mathbb{T}^d))$ solving* (C-CE) *in the sense of Definition 3.2.1, with $u_0 = 0$. Then for every $\varepsilon > 0$ it holds that $\mathbb{M}(T) < \varepsilon$, i.e. $T = 0$.*

Before starting to prove the necessary preliminary lemmas, let us give an idea of the strategy of the proof. The key point is that if we have a 1-current $T_\delta = u_\delta e_t + u_\delta b_\delta^i e_i$ solving (C-CE) with null boundary in $(-\infty, 1) \times \mathbb{T}^d$, supported on $[0, 1) \times \mathbb{T}^d$ and with velocity b_δ a smooth approximation b_δ of b, we can represent T_δ (which can be easily shown to converge to T in mass as $\delta \to 0$) as the push-forward of a time-oriented [1] 1-current whose mass is controlled by the $L^1((0,1) \times \mathbb{T}^d)$ norm of a commutator term r_δ which tends to 0 in $L^1((0,T) \times \mathbb{T}^d)$ as $\delta \to 0$. We stress the fact that the language of currents provides a natural setting in which we can apply the notion of push-forward and bypass the need to use renormalized solutions. This geometric approach may also be of help in treating cases outside of DiPerna-Lions setting. Notice that in Theorem 3.2.1 we prove uniqueness as a consequence of the fact that $\mathbb{M}(T) < \varepsilon$ for every $\varepsilon > 0$ fixed. In principle we could conclude that $T = 0$ even just proving, for example, that the flat norm of T is smaller than ε for every $\varepsilon > 0$, and the flat norm is weaker than the mass norm. It may be that such an estimate can be deduced by weaker estimates on the commutator, which in turn might be obtained even outside of the DiPerna-Lions setting, but this will require further investigations.

Let us start with a simple lemma:

Lemma 3.2.1.
Let T be a normal 1-current such that $\partial T = g$ in $(-\infty, 1) \times \mathbb{T}^d$ with $g \in L^1((0,1) \times \mathbb{T}^d)$, and with $\operatorname{Supp}(T) \subseteq [0,1) \times \mathbb{T}^d$. Then there exists a time-oriented normal 1-current M such that $\partial M = g$ in $(-\infty, 1) \times \mathbb{T}^d$, with support in $[0, 1) \times \mathbb{T}^d$ and satisfying

$$\mathbb{M}(M) \leq \|g\|_{L^1((0,1)\times\mathbb{T}^d)}.$$

Moreover, if ∂T is supported in a set $[0, 1) \times K$ for some K, so is M.

Proof. Consider

$$M(t,x) := -G(t,x) e_t \text{ with } G(t,x) = \begin{cases} \int_0^t g(s,x) ds & \text{if } t > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Given $\xi \in C_c^\infty([0,1) \times \mathbb{T}^d)$:

$$\langle \partial M, \xi \rangle = \langle M, d\xi \rangle = \int_{\mathbb{T}^d} \left(\int_0^1 -G(t,x) \partial_t \xi(t,x) dt \right) dx.$$

[1] We refer to a 1-current completely in the e_t direction as “time-oriented”.

But

$$\int_0^1 -G(t,x)\partial_t\xi(t,x)dt = \overbrace{[-G(t,x)\xi(t,x)]_0^1}^{=0} + \int_0^1 \partial_t G(t,x)\xi(t,x)dt$$
$$= \int_0^1 g(t,x)\xi(t,x)dt,$$

so $\partial G = g$. If g is supported in $[0,1)\times K$, it is clear from the construction that M is supported in $[0,1)\times K$ too, and the mass norm estimate is trivial by construction. □

Let

$$b_\delta(t,\cdot) := \eta_\delta * b(t,\cdot) \text{ and } u_\delta(t,\cdot) := \eta_\delta * u(t,\cdot) \tag{3.2.1}$$

with $\eta_\delta(x) := \delta^{-d}\eta(x/\delta)$ where η is a standard mollifier in the spatial variables only. Consider the current

$$T_\delta := u_\delta e_t + u_\delta b_\delta^i e_i. \tag{3.2.2}$$

As already anticipated, the following commutator estimate will be crucial. Notice that contrary to [70] where only the density u is regularised, we regularise both u and the velocity field b since we will need to use the flow of b_δ (see Lemma 3.2.4). However, as it will be shown in the proof, regularising also the velocity field does not create any problem and we can reduce the proof to the one in [70].

Lemma 3.2.2.
Let $u \in L^\infty(0,1;L^q(\mathbb{T}^d))$ be a solution of the continuity equation with velocity field b and $u_0 = 0$, and let b satisfy (Hyp). *Define u_δ, b_δ as in* (3.2.1). *It holds*

$$\partial_t u_\delta + \operatorname{div}_x(b_\delta u_\delta) = r_\delta, \tag{3.2.3}$$

where $r_\delta := \operatorname{div}_x(b_\delta u_\delta) - \eta_\delta * \operatorname{div}_x(bu) \to 0$ as $\delta \to 0$ strongly in $L^1((0,1)\times\mathbb{T}^d)$. We also point out that $r_\delta = -\partial T_\delta$.

Proof. Let $c := \operatorname{div}_x(b)$ and $c_\delta := \eta_\delta * \operatorname{div}_x(b)$. We can rewrite (3.2.3) as

$$\begin{aligned}\partial_t u_\delta + \operatorname{div}_x(b_\delta u_\delta) &= r_\delta = c_\delta u_\delta + b_\delta\cdot\nabla_x u_\delta - c_\delta u - \eta_\delta*(b\cdot\nabla_x u)\\ &= \underbrace{b\cdot\nabla_x u_\delta - \eta_\delta*(b\cdot\nabla_x u) + cu_\delta - \eta_\delta*(cu)}_{:=r_0^\delta} + \underbrace{[c_\delta u_\delta - cu_\delta]}_{:=r_1^\delta} + \underbrace{[\eta_\delta*(cu) - c_\delta u]}_{:=r_2^\delta} + \underbrace{[(b_\delta - b)\cdot\nabla_x u_\delta]}_{:=r_3^\delta}.\end{aligned}$$

It is trivial to see that the terms r_1^δ, r_2^δ go to 0 in L^1 as $\delta\to 0$, while $r_0^\delta \to 0$ as $\delta \to 0$ by [70, Theorem II.1] (see also Theorem 2.1.3 in Chapter 2). As for r_3^δ, notice that:

$$b_\delta(t,x) - b(t,x) = \int_{\mathbb{T}^d}(b(t,y)-b(t,x))\eta_\delta(x-y)dy \text{ and } \nabla_x u_\delta = \int_{\mathbb{T}^d}\frac{u(t,y)}{\delta}\nabla_x\eta_\delta(x-y)dy.$$

Thus:

$$\begin{aligned}\|r_3^\delta\|_{L^1((0,1)\times\mathbb{T}^d)} &= \int_0^1\int_{\mathbb{T}^d}\left|\left(\int_{\mathbb{T}^d}(b(t,y)-b(t,x))\eta_\delta(x-y)dy\right)\cdot\left(\int_{\mathbb{T}^d}\frac{u(t,y)}{\delta}\nabla_x\eta_\delta(x-y)\right)\right|dxdt\\ &= \int_0^1\int_{\mathbb{T}^d}\left|\left(\int_{\mathbb{T}^d}\frac{b(t,y)-b(t,x)}{\delta}\frac{1}{\delta^d}\eta\left(\frac{x-y}{\delta}\right)dy\right)\cdot\left(\int_{\mathbb{T}^d}u(t,y)\frac{1}{\delta^d}\nabla_x\eta\left(\frac{x-y}{\delta}\right)\right)\right|dxdt.\end{aligned}$$

Changing variables, setting $z = (x-y)/\delta$:

$$\|r_3^\delta\|_{L^1((0,1)\times\mathbb{T}^d)} = \int_0^1 \int_{\mathbb{T}^d} \left| \left(\int_{B_1} \frac{b(t,x-\delta z)-b(t,x)}{\delta} \eta(z)dz \right) \cdot \left(\int_{B_1} u(t,x-\delta z)\nabla_z \eta(z)dz \right) \right| dxdt,$$

with $B_1 \subset \mathbb{T}^d$ the ball of radius 1 centered in 0. Arguing as in Lemma 2.1.1 it is easy to see that, as $\delta \to 0$, the first factor tends to

$$\nabla_x b(t,x) \int_{B_1} z\eta(z)dz = 0$$

since $z\eta(z)$ is odd. As for the second factor it converges to

$$u(t,x) \int_{B_1} \nabla_z \eta(z)dz = 0,$$

since $\nabla_z \eta(z)$ is odd. Thus, $r_3^\delta \to 0$ in $L^1((0,1)\times\mathbb{T}^d)$ as $\delta \to 0$ and the proof is complete.

□

We shall now prove some lemmas regarding the “straightening” procedure for T which will help us make use of Lemma 3.2.1. First of all, we need the following general result on the push-forward of 1-currents.

Lemma 3.2.3.
Let $F : \mathbb{T}^d \to \mathbb{T}^d$ be an invertible C^1 map with C^1 inverse and let $T = f_1 e_1 + \dots f_n e_n$ be a 1-current in $\mathbb{T}^d$ with $f_1, \dots f_n \in L^1(\mathbb{T}^d)$. Then:

$$F_\# T = \left(DF \cdot \frac{T}{\det(DF)} \right) \circ F^{-1}, \tag{3.2.4}$$

with DF denoting the Jacobian matrix of F.

Remark 3.2.3.
This is, in some sense, the vector-valued case of the well-known representation formula for the push-forward of absolutely continuous measures:

$$F_\#(u\mathscr{L}^d) = \left(\frac{u}{\det(DF)} \circ F^{-1} \right) \mathscr{L}^d. \tag{3.2.5}$$

Remark 3.2.4.
We committed a slight abuse of notation in (3.2.4). *In the left-hand side of the equation we view T as a current, while on the right-hand side we treat it as a (column) vector field, multiplying it with a matrix.*

Proof. It is just a computation. Consider $\omega = h^i dx_i$ a 1-form. By definition of push-forward of a current

$$\langle F_\# T, \omega \rangle = \langle T, F^\# \omega \rangle. \tag{3.2.6}$$

Now,

$$F^\# \omega = h^i(F(x)) dF^i = \left(h^i(F(x)) \frac{\partial F^i}{\partial x_j}(x) \right) dx_j, \tag{3.2.7}$$

so

$$\begin{aligned} \langle F_\# T, \omega \rangle &= \int_{\mathbb{T}^d} h^i(F(x)) \frac{\partial F^i}{\partial x_j}(x) f_j(x) dx \\ &= \int_{\mathbb{T}^d} h^i(y) \underbrace{\frac{\partial F^i}{\partial x_j}(F^{-1}(y)) f_j(F^{-1}(y)) |\det(DF^{-1}(y))|}_{=(F_\# T)_i} dy, \end{aligned}$$

where in the last equality we changed variables setting $y = F(x)$. Finally, noticing that $I = D(F \circ F^{-1}(y)) = DF(F^{-1}(y)) \cdot DF^{-1}(y)$ we have

$$\det(DF^{-1}(y)) = \frac{1}{\det(DF(F^{-1}(y)))}$$

by Binet and so we have the thesis. □

Lemma 3.2.4. ***Straightening lemma***
Let

$$T_\delta = u_\delta e_t + u_\delta b^i_\delta e_i \text{ be a 1-current in } (-\infty, 1) \times \mathbb{T}^d$$

and

$$\psi_\delta(t,x) := (t, \Phi_\delta(t,x)),$$

with b_δ, u_δ defined as in (3.2.1) *and Φ_δ the flow of b_δ. It holds that*

$$(\psi_\delta^{-1})_\# T_\delta = \left[\left(\frac{u_\delta}{\det(D_x \Phi_\delta^{-1})} \right) \circ \psi_\delta(t,x) \right] e_t.$$

Proof. By Lemma 3.2.3 the representation (3.2.4) holds with $F = \psi_\delta^{-1}$. We have:

$$D\psi_\delta^{-1}(\psi_\delta(t,x)) = \left[\begin{array}{c|c} 1 & 0 \\ \hline \partial_t \Phi_\delta^{-1}(t, \Phi_\delta(t,x)) & D_x \Phi_\delta^{-1}(t, \Phi_\delta(t,x)) \end{array} \right]$$

and since $\Phi_\delta^{-1}(t, \Phi_\delta(t,x)) = x$ for every (t,x) and $\partial_t \Phi_\delta(t,x) = b_\delta(t, \Phi_\delta(t,x))$:

$$\begin{aligned} 0 &= \frac{d}{dt}(\Phi_\delta^{-1}(t, \Phi_\delta(t,x))) = \partial_t \Phi_\delta^{-1}(t, \Phi_\delta(t,x)) + D_x \Phi_\delta^{-1}(t, \Phi_\delta(t,x)) \partial_t \Phi_\delta(t,x) = \\ &= \partial_t \Phi_\delta^{-1}(t, \Phi_\delta(t,x)) + D_x \Phi_\delta^{-1}(t, \Phi_\delta(t,x)) b_\delta(t, \Phi_\delta(t,x)). \end{aligned} \tag{3.2.8}$$

Then we can rewrite:

$$D\psi_\delta^{-1}(\psi_\delta(t,x)) = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & D_x \Phi_\delta^{-1}(t, \Phi_\delta(t,x)) \end{array} \right] \left[\begin{array}{c|c} 1 & 0 \\ \hline -b_\delta(t, \Phi_\delta(t,x)) & I \end{array} \right]$$

and by Lemma 3.2.3:

$$\begin{aligned} (\psi_\delta^{-1})_\#(T_\delta) &= \left(\frac{1}{\det(D_x \Phi_\delta^{-1})} \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & D_x \Phi_\delta^{-1} \end{array} \right] \left[\begin{array}{c|c} 1 & 0 \\ \hline -b_\delta & I \end{array} \right] \left[\begin{array}{c} u_\delta \\ u_\delta b_\delta \end{array} \right] \right) \underbrace{(t, \Phi_\delta(t,x))}_{=\psi_\delta(t,x)} = \\ &= \left[\left(\frac{u_\delta}{\det(D_x \Phi_\delta^{-1})} \right) \circ \psi_\delta(t,x) \right] e_t. \end{aligned}$$

□

We are now ready to conclude.

Proof of Theorem 3.2.1. Fix $\delta > 0$ and consider b_δ, u_δ as in (3.2.1) and $T_\delta := u_\delta e_t + u_\delta b^i_\delta e_i$. We have:

$$T = (T - T_\delta) + T_\delta. \tag{3.2.9}$$

By construction $\mathbb{M}(T - T_\delta) \to 0$ as $\delta \to 0$. So we only need to prove that

$$\mathbb{M}(T_\delta) \to 0 \text{ as } \delta \to 0.$$

In order to do that, we consider the push-forward of T_δ via ψ_δ^{-1}, where we recall that $\psi_\delta(t,x) = (t, \Phi_\delta(t,x))$, and by Lemma 3.2.4 we have

$$(\psi_\delta^{-1})_\# T_\delta = \left[\left(\frac{u_\delta}{\det(D_x \Phi_\delta^{-1})} \right) \circ (t, \Phi_\delta(t,x)) \right] e_t.$$

Consider

$$\tilde r_\delta := \partial(\psi_\delta^{-1})_\# T_\delta = (\psi_\delta^{-1})_\# \partial T_\delta = -(\psi_\delta^{-1})_\# r_\delta,$$

with r_δ defined as in Lemma 3.2.2. Using the representation formula (3.2.5) for push-forward measures we can easily see that $\|\tilde r_\delta\|_{L^1((0,1)\times\mathbb{T}^d)} = \|r_\delta\|_{L^1((0,1)\times\mathbb{T}^d)}$, indeed for every $t \in (0,1)$ fixed:

$$\begin{aligned} \int_{\mathbb{T}^d} |\tilde r_\delta|(t,x)dx &= \int_{\mathbb{T}^d} \frac{|r_\delta|(t,\Phi_\delta(t,x))}{\det(D_x\Phi_\delta^{-1}(t,\Phi_\delta(t,x)))} dx = \\ &= \int_{\mathbb{T}^d} \frac{|r_\delta|(t,y)}{\det(D_x\Phi_\delta^{-1}(t,y))\det(D_x\Phi_\delta(t,\Phi_\delta^{-1}(t,y)))} dy = \int_{\mathbb{T}^d} |r_\delta|(t,y)dy, \end{aligned} \tag{3.2.10}$$

where we used the change of variables $y = \Phi_\delta(t,x)$ and the fact that

$$\det(D_x\Phi_\delta^{-1}(t,y))\det(D_x\Phi_\delta(t,\Phi_\delta^{-1}(t,y))) = \det(\underbrace{D_x\Phi_\delta^{-1}(t,y)\cdot D_x\Phi_\delta(t,\Phi_\delta^{-1}(t,y))}_{=I}) \equiv 1.$$

By Lemma 3.2.1, $\exists P_\delta$ a time-oriented 1-current with mass controlled by $\|\tilde r_\delta\|_{L^1((0,1)\times\mathbb{T}^d)} = \|r_\delta\|_{L^1((0,1)\times\mathbb{T}^d)}$ and such that $\partial P_\delta = \tilde r_\delta$. By construction $(\psi_\delta)^{-1}_\# T_\delta - P_\delta$ has null boundary in $(-\infty,1)\times\mathbb{T}^d$ and it is completely time-oriented, so it is zero. Indeed let $h(t,x)$ be such that $(\psi_\delta)^{-1}_\# T_\delta - P_\delta = h(t,x)e_t$. The null boundary condition can be written in this case as:

$$\partial_t h = 0 \text{ as a distribution} \implies h(\cdot,x) \text{ constant } \forall x \implies h \equiv 0,$$

because the current has compact support in time. We thus have:

$$T_\delta = (\psi_\delta)_\# P_\delta \implies \mathbb{M}(T_\delta) = \mathbb{M}((\psi_\delta)_\# P_\delta).$$

We can now calculate $\mathbb{M}((\psi_\delta)_\# P_\delta)$. Consider:

- A 1-form $\omega(t,x) = \tau(t,x)dt + \xi^j(t,x)dx_j$ with $\tau, \xi^j \in C_c^\infty((-\infty,1)\times\mathbb{T}^d)$ such that $|\tau|, |\xi^j| \le 1$ for every j;
- A function $f \in L^1((0,1)\times\mathbb{T}^d)$ such that $P_\delta(t,x) = f(t,x)dt$. We refer the reader to Lemma 3.2.1 for the precise definition of f, for which it holds $\|f\|_{L^1((0,1)\times\mathbb{T}^d)} \le \|\tilde r_\delta\|_{L^1((0,1)\times\mathbb{T}^d)} = \|r_\delta\|_{L^1((0,1)\times\mathbb{T}^d)}$ by construction.

Then:

$$(\psi_\delta)_\# P_\delta(\omega) = P_\delta((\psi_\delta)^\# \omega) = P_\delta(\tau(\psi_\delta(t,x))d\psi_\delta^t + \xi^j(\psi_\delta(t,x))d\psi_\delta^j),$$

where with ψ_δ^t we denote the component of ψ_δ in the direction e_t and with ψ_δ^j the components of ψ_δ in the directions e_j. Since P_δ is completely time-oriented, only the partial derivatives with respect to the time variable will play a role (this is the so-called *tangential differential*), i.e.

$$\begin{aligned}
&P_\delta(\tau(\psi_\delta(t,x))d\psi_\delta^t + \xi^j(\psi_\delta(t,x))d\psi_\delta^j) \\
&= P_\delta\left(\left(\tau(\psi_\delta(t,x)) + \sum_{j=1}^d b_\delta^j(\psi_\delta(t,x))\xi^j(\psi_\delta(t,x))\right)dt\right) \\
&= \int_0^1\int_{\mathbb{T}^d}\left(\tau(\psi_\delta(t,x)) + \sum_{j=1}^d b_\delta^j(\psi_\delta(t,x))\xi^j(\psi_\delta(t,x))\right)f(t,x)dxdt,
\end{aligned} \tag{3.2.11}$$

where we used that $\partial_t\psi_\delta^j(t,x) = b^j(t,\Phi_\delta(t,x)) = b^j(\psi_\delta(t,x))$ for every $j = 1,\dots,d$. But since $b \in L^\infty((0,1)\times\mathbb{T}^d)$ and by assumption $|\tau|, |\xi^j| \le 1$ for every $j = 1,\dots,d$ we have that

$$\mathbb{M}(T_\delta) = \mathbb{M}((\psi_\delta)_\# P_\delta) \lesssim \mathbb{M}(P_\delta) = \|r_\delta\|_{L^1((0,1)\times\mathbb{T}^d)} \to 0 \text{ as } \delta \to 0, \tag{3.2.12}$$

and we conclude since, by choosing δ small enough, we can make the mass of T arbitrarily small.

□

Remark 3.2.5. ***BV velocity fields***
The previous result can actually be extended to the case of a velocity field $b \in [L^1(0,1;BV(\mathbb{T}^d)) \cap L^\infty((0,1)\times\mathbb{T}^d)]^d$. The only difference would be in the proof of Lemma 3.2.2 where the choice of a standard mollifier η will not be enough to ensure that the $L^1((0,1)\times\mathbb{T}^d)$ norm of the commutator r_δ is small. This can be fixed by employing ad hoc local anisotropic mollifiers, although the procedure is much more involved. For the details, see [7].

Remark 3.2.6. ***The role of the divergence***
Notice that we have never used the bound on $\operatorname{div}_x(b)^-$. The reason is that, for proving the uniqueness, this is not needed. However it is needed for proving existence of solutions of solutions of the continuity equation, see Proposition 2.1.1.

Remark 3.2.7. ***The transport equation and the non-homogeneous continuity equation***
It is obvious that, since the issue lies in proving the uniqueness of solutions and the continuity equation is linear, the exact same approach also proves the uniqueness of

$$\begin{cases} \partial_t u + \operatorname{div}_x(bu) = f & \text{in } (0,1)\times\mathbb{T}^d \\ u(0,\cdot) = u_0 & \text{in } \mathbb{T}^d \end{cases} \tag{3.2.13}$$

with $f \in L^1((0,1)\times\mathbb{T}^d)$. With a little additional work, this approach can also establish the uniqueness of the solution of the transport equation, namely

$$\begin{cases} \partial_t\rho + b\cdot\nabla_x\rho = 0 & \text{in } (0,1)\times\mathbb{T}^d \\ \rho(0,\cdot) = \rho_0 & \text{in } \mathbb{T}^d \end{cases}.$$

In order to prove it, we will prove the uniqueness of solutions of (3.2.13) *when in the right-hand side we have a term cu with $c \in L^1(0,T;L^p(\mathbb{T}^d))$ and $u_0 = 0$ (so that the equation becomes essentially like* (DP-L), *see Chapter 2). Then, we have uniqueness of the solutions of the transport equation simply choosing $c = \operatorname{div}_x(b)$. We can consider this right-hand side as a non-zero boundary term. Namely, we need to prove the uniqueness of solutions for* (C-CE) *assuming that $T = ue_t + ub^ie_i$ and*

$$\partial T = -cu. \tag{3.2.14}$$

By repeating the same arguments of the null boundary case, when we regularise the current to have $T_\delta = u_\delta e_t + u_\delta b_\delta^i e_i$, we can consider to have, on the right-hand side of (3.2.14), *$-\eta_\delta * (cu) - r_\delta$, with r_δ being the commutator term defined in Lemma 3.2.2. Following the same proof we end up having, instead of* (3.2.12), *the estimate*

$$\mathbb{M}(T) \le K\|cu\|_{L^1((0,1)\times\mathbb{T}^d)},$$

for some constant $K > 0$ which only depends on the dimension d and on $\|b\|_{L^\infty}$ (check (3.2.11) *to see where $\|b\|_\infty$ comes into play). But the mass of T is precisely $\|u\|_{L^1((0,1)\times\mathbb{T}^d)} + \|bu\|_{L^1((0,1)\times\mathbb{T}^d)}$, so we have*

$$\|u\|_{L^1((0,1)\times\mathbb{T}^d)} \le \|u\|_{L^1((0,1)\times\mathbb{T}^d)} + \|bu\|_{L^1((0,1)\times\mathbb{T}^d)} \le K\|cu\|_{L^1((0,1)\times\mathbb{T}^d)}.$$

We first prove that the uniqueness holds when $|c| < \varepsilon$ for some $\varepsilon > 0$. Indeed, choosing $\varepsilon = 1/(2K)$ we immediately get the thesis. If $c \in L^\infty((0,T)\times\mathbb{T}^d)$ we just have to notice that for every $0 < \lambda < 1$ we have that the functions

$$u_\lambda(t,x) := u(\lambda t, \lambda x) \text{ and } b_\lambda := b(\lambda t, \lambda x)$$

solve

$$\begin{cases} \partial_t u_\lambda + \operatorname{div}_x(u_\lambda b_\lambda) = c\lambda u_\lambda \\ u_\lambda(0,\cdot) = 0 \end{cases} \quad \text{in } [0,1/\lambda) \times \frac{\mathbb{T}^d}{\lambda}.$$

Then, for any given (bounded) c we just have to choose λ small enough so that $|\lambda c| \le 1/(2K)$ (recall that K only depends on d and $\|b_\lambda\|_{L^\infty} = \|b\|_{L^\infty}$) to have $u_\lambda = 0$ in $[0,1/\lambda) \times \frac{\mathbb{T}^d}{\lambda}$, which implies $u = 0$. The general case assuming $c \in L^1(0,T;L^p(\mathbb{T}^d))$ can be achieved by a standard approximation argument.

Chapter 4

Quantitative estimates: a direct approach to regular Lagrangian flows

In this chapter, we shall give *quantitative* stability estimates on regular Lagrangian flows relative to different vector fields, and we shall use such Lagrangian estimates to prove analogous results at the level of the continuity and transport equations. The question is the following: we know, by stability results in previous chapters, that regular Lagrangian flows are stable with respect to their vector fields. Namely, we saw in Theorem 2.2.2 [1] that if a sequence $\{b_n\}_{n\in\mathbb{N}}$ converges strongly in $L^1(I \times \mathbb{T}^d)$ to a bounded vector field b with bounded divergence and if $\exists C > 0$ such that

$$\|b_n\|_{L^\infty(I\times\mathbb{T}^d))} + \|\operatorname{div}_x(b_n)\|_{L^\infty(I\times\mathbb{T}^d))} \leq C \quad \text{for every } n,$$

then the regular Lagrangian flows X_n of b_n converge strongly in $L^1(I \times \mathbb{T}^d)$ to the regular Lagrangian flow X of b. Can we give a quantitative estimate on how close X_n is to X based on how close b_n is to b?

In general, the answer is (for the moment at least) no. The reason is that the stability properties for the regular Lagrangian flow are proved, both in [70] and in [7], by methods (the renormalization technique and the superposition principle, respectively) which are inherently non-quantitative. While these techniques are very powerful in showing well-posedness of the regular Lagrangian flow, they do not seem apt at providing quantitative estimates. The same holds for the current-based approach of Chapter 3. The seminal work in this direction is [58], where the authors manage to prove quantitative estimates. Their argument, however, only works with vector fields with spatial $W^{1,p}$ regularity with $p > 1$ because, as we shall see, the argument crucially relies on the boundedness of the maximal function from L^p to L^p, which only holds if $p > 1$.

Using the results from [58], in [144] quantitative estimates for the solution of (CE) are derived. We also remark that in [58] other properties of regular Lagrangian flows [2] are proved, namely:

- The Lipschitz regularity (in the spatial variables) outside a set of measure $\varepsilon > 0$ arbitrarily small, with a Lipschitz constant $L = L(\varepsilon) \to +\infty$ as $\varepsilon \to 0$;
- the regular Lagrangian flow preserves some regularity [3] of the initial datum u_0 of (CE);
- a mixing conjecture proposed in [33] is answered in the positive.

Since the estimates in [58] and the techniques of the proofs will be the basis for the main results of the next chapters, we give a proof of the results of both [58] and [144]. We shall also prove (since we will use it repeatedly later) the Lipschitz regularity result for the regular Lagrangian flow mentioned above. The proofs are given in the case of a bounded vector field in $\mathbb{T}^d$, but they hold even in $\mathbb{R}^d$ with b unbounded if we localise the estimates and provided b satisfies some additional growth assumptions. We refer the interested reader to [58] for the details.

[1] We also recall Bressan's compactness conjecture [32], which has now been proved. See Theorem 2.1.7.

[2] A previous result on this topic was obtained in [14], where the authors proved approximate differentiability of the regular Lagrangian flow. This property is proved again in [58] in a completely different way.

[3] Namely, the regularity which is preserved is a quite weak one called Lipexp_p regularity, and it is a sort of weak Lipschitz regularity, check [58] for the details.

4.1 The key lemmas

In this brief section we will simply state the three main technical lemmas we will need to use in the following. Despite being only three lemmas, they are so fundamental in order to get quantitative estimates that listing them in a section of their own seems justified. We point out that in this chapter we will only use Lemma 4.1.1 and Lemma 4.1.2. As for Lemma 4.1.3, the asymmetric estimate it provides will be used in the next chapters. We will state the results in $\mathbb{R}^d$ to be consistent with the references, but they also hold in $\mathbb{T}^d$.

Definition 4.1.1. ***Local maximal function***
Let μ be a (vector-valued) locally finite measure in $\mathbb{R}^d$ and let $|\mu|$ be its total variation. For every $\lambda > 0$ we define the local maximal function *of μ as*

$$M_\lambda \mu(x) := \sup_{0<r<\lambda} \frac{|\mu|(B_r(x))}{|B_r(x)|} = \sup_{0<r<\lambda} \fint_{B_r(x)} d|\mu|(y), \text{ for } x \in \mathbb{R}^d.$$

When $\mu = f\mathscr{L}^d$, where f is a L^1_{loc} function, we will use the notation $M_\lambda f$ to denote $M_\lambda \mu$. If $\lambda = +\infty$ and μ is a finite measure in $\mathbb{R}^d$, we denote by

$$M\mu(x) := \sup_{r>0} \frac{|\mu|(B_r(x))}{|B_r(x)|} = \sup_{r>0} \fint_{B_r(x)} d|\mu|(y), \text{ for } x \in \mathbb{R}^d$$

the maximal function *of μ. As we did for the local maximal function, we use the notation Mf if $\mu = f\mathscr{L}^d$ with $f \in L^1$.*

Lemma 4.1.1. ***Key lemma 1: $L^p - L^p$ boundedness of the maximal function, [149]***
Let $f \in L^p_{loc}(\mathbb{R}^d)$ with $p > 1$, and let $\lambda, r > 0$. Then:

$$\int_{B_r(0)} (M_\lambda f(y))^p dy \le c_{d,p} \int_{B_{r+\lambda}(0)} |f(y)|^p dy,$$

with $c_{d,p}$ a constant depending on d and p.

Lemma 4.1.2. ***Key lemma 2: symmetric pointwise inequality for BV functions, [149]***
Let $f \in BV_{loc}(\mathbb{R}^d)$. Then there exists a $\mathscr{L}^d$-negligible set $N \subset \mathbb{R}^d$ such that

$$|f(x) - f(y)| \le c_d |x-y|[M_\lambda |Df|(x) + M_\lambda |Df|(y)] \text{ for every } x, y \in \mathbb{R}^d \setminus N,$$

where M_λ is the local maximal function with $\lambda \ge |x-y|$ and c_d is a constant which only depends on d.

Lemma 4.1.3. ***Key lemma 3: asymmetric pointwise inequality for Sobolev functions, [43, Lemma 5.1]***
Let $f \in W^{1,p}_{loc}(\mathbb{R}^d)$ with $p > d$. Then there exists a $\mathscr{L}^d$-negligible set $N \subset \mathbb{R}^d$ such that

$$|f(x) - f(y)| \le c_{d,p} |x-y|[M_\lambda(|Df|^p)(x)]^{1/p} \text{ for every } x, y \in \mathbb{R}^d \setminus N,$$

where M_λ is the local maximal function with $\lambda \ge |x-y|$ and $c_{d,p}$ is a constant which only depends on d and p.

4.2 Quantitative stability estimates for regular Lagrangian flows

We start this section with a theorem which will be useful in proving the spatial Lipschitz regularity of the regular Lagrangian flow outside a set of small measure. As usual, we will state the results in $\mathbb{T}^d$ rather than $\mathbb{R}^d$ for consistency with results in other chapters, but the case in $\mathbb{R}^d$ is basically identical.

Theorem 4.2.1. ***[58]***
Let $b \in [L^1(0,T; W^{1,p}(\mathbb{T}^d))]^d$ be a bounded vector field with $p > 1$ and let X be the regular Lagrangian flow associated to b. Let L be the compressibility constant of X, as defined in Definition 2.2.4. For every $p > 1$ we define the following integral quantity:

$$A_p(R, X) := \left[\int_{B_R(0)} \left(\sup_{0 \le t \le T} \sup_{0<r<2R} \fint_{B_r(x)} \log \left(1 + \frac{|X(t,x) - X(t,y)|}{r} \right) dy \right)^p dx \right]^{1/p}. \tag{4.2.1}$$

Then we have:

$$A_p(R, X) \le C(p, d, R, L, \|D_x b\|_{L^1(0,T;L^p(\mathbb{T}^d))}).$$

Since we assumed to work in a torus, if R is big enough we have that $\mathbb{T}^d = B_R(x)$ for every $x \in \mathbb{T}^d$, so we can define

$$A_p(X) := \inf_{\{R | \mathbb{T}^d \subset B_R(x) \ \forall x \in \mathbb{T}^d\}} A_p(R, X).$$

Before proving Theorem 4.2.1, we show its application in the proof of the following proposition.

Proposition 4.2.1. ***Lipschitz regularity of the regular Lagrangian flow outside a small set, [58]***
Let $X : [0,T] \times \mathbb{T}^d \to \mathbb{T}^d$ be a map. Then, for every $\varepsilon > 0$ and every $R > 0$ we can find a set $K \subset B_R(0)$ such that $|B_R(0) \setminus K| \le \varepsilon$ and for any $0 \le t \le T$ it holds [4]

$$\mathrm{Lip}(X(t,\cdot)|_K) \le \exp\left(\frac{2A_p(R,X)}{\varepsilon^{1/p}} \right).$$

Proof. Fix $\varepsilon > 0$ and $R > 0$. Assuming $A_p(R, X)$ to be finite (otherwise there is nothing to prove) we consider

$$M := \frac{A_p(R,X)}{\varepsilon^{1/p}}.$$

By Chebyshev's inequality we have that for every $t > 0$ and for every $f \in L^p(B_R(0))$:

$$\mathscr{L}^d(\{x \in B_R(0) \text{ such that } |f(x)| \ge t\}) \le \frac{1}{t^p} \int_{B_R(0)} |f(y)|^p dy.$$

Choosing

$$t = M \ \text{ and } \ f(x) = \sup_{0 \le t \le T} \sup_{0<r<2R} \fint_{B_r(x)} \log \left(1 + \frac{|X(t,x) - X(t,y)|}{r} \right) dy,$$

we have that

$$f(x) \le M$$

for every $x \in K \subset B_R(0)$ with $|B_R(0) \setminus K| \le \varepsilon$. Now, fix $x, y \in K$. It holds $|x - y| < 2R$ and by setting $r = |x - y|$ we have:

$$\begin{aligned}
&\log \left(1 + \frac{|X(t,x) - X(t,y)|}{r} \right) = \fint_{B_r(x) \cap B_r(y)} \log \left(1 + \frac{|X(t,x) - X(t,y)|}{r} \right) dz \\
&\le \fint_{B_r(x) \cap B_r(y)} \log \left(1 + \frac{|X(t,x) - X(t,z)|}{r} + \frac{|X(t,y) - X(t,z)|}{r} \right) dz,
\end{aligned}$$

where we simply used the monotonicity of the function $\log(1 + x)$. By the elementary inequality

$$\log(1 + x) + \log(1 + y) = \log(1 + x + y + xy) \ge \log(1 + x + y) \ \text{ for } x, y \ge 0,$$

we have:

[4] In [58] the factor 2 inside the argument of the exponential is substituted by a dimension-dependent constant C_d.

$$
\begin{aligned}
&\fint_{B_r(x)\cap B_r(y)} \log\left(1+\frac{|X(t,x)-X(t,z)|}{r}+\frac{|X(t,y)-X(t,z)|}{r}\right)dz\\
&\le \fint_{B_r(x)\cap B_r(y)} \log\left(1+\frac{|X(t,x)-X(t,z)|}{r}\right)dz+\fint_{B_r(x)\cap B_r(y)} \log\left(1++\frac{|X(t,y)-X(t,z)|}{r}\right)dz\\
&\le \fint_{B_r(x)} \log\left(1+\frac{|X(t,x)-X(t,z)|}{r}\right)dz+\fint_{B_r(y)} \log\left(1+\frac{|X(t,x)-X(t,z)|}{r}\right)dz\\
&\le 2M=\frac{2A_p(R,X)}{\varepsilon^{1/p}}.
\end{aligned}
$$

These inequalities hold uniformly in time, so for every $t \in [0,T]$:

$$
\log\left(1+\frac{|X(t,x)-X(t,y)|}{r}\right)\le\frac{2A_p(R,X)}{\varepsilon^{1/p}} \implies |X(t,x)-X(t,y)|\le\exp\left(\frac{2A_p(R,X)}{\varepsilon^{1/p}}\right)|x-y|
$$

for every $x, y \in K$, and the proof is complete. □

We now give the proof of Theorem 4.2.1.

Proof of Theorem 4.2.1. For $0 \le t \le T$, $0 < r < 2R$ and $x \in B_R(0)$ we define the quantity

$$
Q(t,x,r) := \fint_{B_r(x)} \log\left(1+\frac{X(t,x)-X(t,y)}{r}\right)dy.
$$

By Definition 2.2.4 we have:

$$
\begin{aligned}
\frac{d}{dt}Q(t,x,r) &\le \fint_{B_r(x)}\left|\frac{dX}{dt}(t,x)-\frac{dX}{dt}(t,y)\right|(|X(t,x)-X(t,y)|+r)^{-1} \\
&= \fint_{B_r(x)} \frac{|b(t,X(t,x))-b(t,X(t,y))|}{|X(t,x)-X(t,y)|+r}dy.
\end{aligned}
\tag{4.2.2}
$$

We choose $\bar R = 4R + 2T\|b\|_{L^\infty}$, so that $|X(t,x) - X(t,y)| < \bar R$ for every $x, y \in B_R(0)$ and for every $t \in [0,T]$. By Lemma 4.1.2 we have:

$$
\begin{aligned}
\frac{dQ}{dt}(t,x,r) &\le c_d \fint_{B_r(x)} \left[M_{\bar R}|Db|(t,X(t,x))+M_{\bar R}|Db|(t,X(t,y))\right]\frac{|X(t,x)-X(t,y)|}{|X(t,x)-X(t,y)|+r}dy \\
&\le c_d M_{\bar R}|Db|(t,X(t,x))+c_d\fint_{B_r(x)} M_{\bar R}|Db|(t,X(t,y))dy.
\end{aligned}
\tag{4.2.3}
$$

Integrating in time up to time $\bar t \in [0,T]$ and taking the supremum over $r \in (0,2R)$ we get:

$$
\begin{aligned}
\sup_{0<r<2R} Q(\bar t,x,r) &\le C + c_d\int_0^{\bar t} M_{\bar R}|Db|(t,X(t,x))dt + c_d\sup_{0<r<2R}\int_0^{\bar t}\fint_{B_r(x)} M_{\bar R}|Db|(t,X(t,y))dydt\\
&\le C + c_d\int_0^{T} M_{\bar R}|Db|(t,X(t,x))dt + c_d\sup_{0<r<2R}\int_0^{T}\fint_{B_r(x)} M_{\bar R}|Db|(t,X(t,y))dydt\\
&\le C + c_d\int_0^{T} M_{\bar R}|Db|(t,X(t,x))dt + c_d\int_0^{T}\sup_{0<r<2R}\fint_{B_r(x)} M_{\bar R}|Db|(t,X(t,y))dydt,
\end{aligned}
$$

where we exchanged the order of the time derivative and the supremum and we used that, for every $r > 0$ and $x \in T^d$, we have $Q(0,x,r) \le \log(2) = C$. Thus:

$$\sup_{0\le \bar t\le T}\sup_{0<r<2R} Q(\bar t,x,r)\le C+c_d\int_0^T M_{\bar R}|Db|(t,X(t,x))dt+c_d\int_0^T \sup_{0<r<2R}\fint_{B_r(x)} M_{\bar R}|Db|(t,X(t,y))dydt.$$

Taking the L^p norm over $B_R(0)$ we get

$$\begin{aligned}
&A_p(R,X)\\
&\le C_{p,R}+c_d\left\|\int_0^T M_{\bar R}|Db|(t,X(t,x))dt\right\|_{L^p(B_R(0))}+c_d\left\|\int_0^T \sup_{0<r<2R}\fint_{B_r(x)} M_{\bar R}|Db|(t,X(t,y))dydt\right\|_{L^p(B_R(0))}\\
&\le C_{p,R}+c_d\int_0^T \left\|M_{\bar R}|Db|(t,X(t,x))\right\|_{L^p(B_R(0))}dt+c_d\int_0^T \left\|\sup_{0<r<2R}\fint_{B_r(x)} M_{\bar R}|Db|(t,X(t,y))dy\right\|_{L^p(B_R(0))}dt\\
&= C_{p,R}+c_d\int_0^T \left\|M_{\bar R}|Db|(t,X(t,x))\right\|_{L^p(B_R(0))}dt+c_d\int_0^T \left\|M_{2R}(M_{\bar R}|Db|(t,X(t,\cdot)))(x)\right\|_{L^p(B_R(0))}dt.
\end{aligned}$$

Using the bounded compressibility L of the regular Lagrangian flow we readily obtain:

$$\begin{aligned}
A_p(R,X)\le C_{p,R}&+c_dL^{1/p}\int_0^T \left\|M_{\bar R}|Db|(t,\cdot)\right\|_{L^p(B_{R+T\|b\|_\infty}(0))}dt\\
&+c_dL^{1/p}\int_0^T \left\|M_{2R}M_{\bar R}|Db|(t,\cdot)\right\|_{L^p(B_{R+T\|b\|_\infty}(0))}dt,
\end{aligned}$$

and we conclude applying Lemma 4.1.1 (twice on the last addendum).

□

Before giving a stability estimate, we will show how using Proposition 4.2.1 we can easily prove in a quite direct way the existence and compactness of the regular Lagrangian flow, whereas in Theorem 2.2.2 the superposition principle and the well-posedness of (μ-CE) was needed. We need a general real analysis lemma whose proof we omit:

Lemma 4.2.1. ***[58]***
Let $\Omega\subset\mathbb{R}^n$ be a bounded Borel set and let $\{f_h\}_{h\in\mathbb{N}}$ be a sequence of maps $f_h:\mathbb{R}^n\to\mathbb{R}^m$. Suppose that for every $\delta>0$ there is a positive constant $M_\delta<+\infty$ and, for every h, a set $B_{h,\delta}\subset\Omega$ with $|\Omega\setminus B_{h,\delta}|\le\delta$ such that

$$\|f_h\|_{L^\infty(B_{h,\delta})}\le M_\delta \text{ and } \mathrm{Lip}(f_h|_{B_{h,\delta}})\le M_\delta.$$

Then the sequence $\{f_h\}_{h\in\mathbb{N}}$ is precompact in measure, i.e. there is a subsequence $\{f_{h_n}\}_{h_n\in\mathbb{N}}$ such that $f_{h_n}\to f$ as $h_n\to+\infty$ in measure. That is, for every $\varepsilon>0$

$$\mathscr{L}^n(\{x\in\Omega \text{ such that } |f_{h_n}(x)-f(x)|>\varepsilon\})\to 0 \text{ as } h_n\to+\infty.$$

We are now ready to prove, in a direct fashion, the compactness and the existence of regular Lagrangian flows.

Corollary 4.2.1. ***A direct proof of compactness of regular Lagrangian flows, [58]***
Let $\{b_h\}_{h\in\mathbb{N}}$ be a sequence of vector fields equibounded in $[L^\infty([0,T]\times\mathbb{T}^d)\cap L^1(0,T;W^{1,p}(\mathbb{T}^d))]^d$ for some $p>1$. For each h let X_h be a regular Lagrangian flow associated to b_h and let L_h be the compressibility constant of the flow. Assume that the $\{L_h\}_{h\in\mathbb{N}}$ is equibounded by a constant $L>0$. Then the sequence $\{X_h\}_{h\in\mathbb{N}}$ is strongly precompact in $L^1([0,T]\times\mathbb{T}^d)$. Moreover, if $b_h\to b\in[L^1(0,T;W^{1,p}(\mathbb{T}^d))]^d$, we can extract subsequences $\{b_{h_n}\},\{X_{h_n}\}$ such that

$$b_{h_n}\to b \text{ and } X_{h_n}\to X \text{ in } L^1([0,T]\times\mathbb{T}^d) \text{ as } h_n\to+\infty,$$

and X is a regular Lagrangian flow of b with compressibility bounded by L.

Proof. Fix $\delta > 0$ and notice that $\{\|X_h\|\}_{h\in\mathbb{N}}$ is equibounded, simply because $\mathbb{T}^d$ is a compact metric space. By Proposition 4.2.1, for every h we can find a Borel set $K_{h,\delta}$ such that

$$|\mathbb{T}^d \setminus K_{h,\delta}| \leq \delta \text{ and } \mathrm{Lip}(X_h(t,\cdot)|_{K_{h,\delta}}) \leq \exp\left(\frac{CA_p(X_h)}{\delta^{1/p}}\right) \text{ for every } t \in [0,T].$$

By the estimates in Theorem 4.2.1 and our assumptions on $\{b_h\}_{h\in\mathbb{N}}, \{L_h\}_{h\in\mathbb{N}}$, we have that $A_p(X_h)$ is equibounded in h , and so $\mathrm{Lip}(X_h(t,\cdot)|_{K_{h,\delta}})$ is equibounded for any fixed δ. Thus, we can apply Lemma 4.2.1 to get convergence in measure of $\{X_h\}_{h\in\mathbb{N}}$. Up to a subsequence, convergence in measure implies convergence almost everywhere, and since $\{X_h\}_{h\in\mathbb{N}}$ is also equibounded in $L^\infty([0,T]\times\mathbb{T}^d)$ (and consequently also on $L^1([0,T]\times\mathbb{T}^d)$) we conclude the convergence in L^1 by dominated convergence. If $b_h \to b \in [L^1(0,T;W^{1,p}(\mathbb{T}^d))]^d$ we can assume that, up to a subsequence which we will not rename:

$$X_h \to X \text{ and } b_h \to b \text{ in } L^1([0,T]\times\mathbb{T}^d) \text{ as } h \to +\infty.$$

In order to conclude, we just need to show that X is a regular Lagrangian flow with respect to b. Assume that condition 1. of Definition 2.2.4 is not satisfied. Then we would have a set A of positive measure such that, for every $x \in A$, the map $t \mapsto X(t,x)$ is not an absolutely continuous solution of $\partial_t X(t,x) = b(t,X(t,x))$ with $X(0,x) = x$, i.e.

$$X(t,x) \neq x + \int_0^t b(s,X(s,x))ds \text{ for every } x \in A \text{ for some } t \in [0,T].$$

Moreover, we can assume without loss of generality that for every $x \in A$ the above inequality holds in a set of times of positive measure, otherwise it would be enough to redefine X in a $\mathscr{L}^{1+d}$-null set. We thus have:

$$\int_0^T \int_A \left| X(t,x) - x + \int_0^t b(s,X(s,x))ds \right| dxdt > 0,$$

but at the same time it is easy to see that

$$0 = \int_0^T \int_A \left| X_h(t,x) - x + \int_0^t b_h(s,X_h(s,x))ds \right| dxdt \to \int_0^T \int_A \left| X(t,x) - x + \int_0^t b(s,X(s,x))ds \right| dxdt$$

as $h \to 0$, a contradiction. The compressibility bound is proved (up to extracting another subsequence) by the tightness of the measures $\{\mathscr{L}^1_t \otimes X_h(t,\cdot)_\# \mathscr{L}^d_x\}_{h\in\mathbb{N}}$ and by Prokhorov's compactness theorem. □

Corollary 4.2.2. ***A direct proof of existence of regular Lagrangian flows, [58]***
Let $b \in [L^1(0,T;W^{1,p}(\mathbb{T}^d))]^d$ be a bounded vector field with $p > 1$ and such that $\mathrm{div}_x(b)^- \in L^1(0,T;L^\infty(\mathbb{T}^d))$. Then there exists a regular Lagrangian flow associated to b.

Proof. It follows easily by Corollary 4.2.1. Consider $\{b_\delta\}_{\delta>0}$ a sequence of regularised vector fields by convolution and consider $\{X_\delta\}_{\delta>0}$ the respective (classical) flows. It is easy to check that both $\{b_\delta\}_{\delta>0}$ and $\{X_\delta\}_{\delta>0}$ satisfy the assumptions of Corollary 4.2.1, in particular $b_\delta \to b$ as $\delta \to 0$ in $[L^1(0,T;W^{1,p}(\mathbb{T}^d)]^d$ and we can uniformly estimate the compressibility constants L_h of X_h (see Proposition 1.1.5):

$$L_\delta \leq \exp\left(\int_0^T \|\,\mathrm{div}_x(b_\delta)^-(t,\cdot)\|_{L^\infty} dt\right) \leq \exp\left(\int_0^T \|\,\mathrm{div}_x(b)^-(t,\cdot)\|_{L^\infty} dt\right),$$

where in the last inequality we used Young's convolution inequality. Applying Corollary 4.2.1, up to a subsequence, we find that $X_\delta \to X$, with X a regular Lagrangian flow for b. □

We now prove the second main result of this section: a quantitative stability inequality between regular Lagrangian flows. We slightly modify the argument of [58] to get a L^p estimate and a time dependent bound, since we will need it in the next chapters.

Proposition 4.2.2. ***L^p stability estimate for regular Lagrangian flows, [58, 54]***
Let b and $\tilde{b}$ be bounded vector fields belonging to $[L^1(0,T;W^{1,p}(\mathbb{T}^d))]^d$ for some $p>1$. Let X and $\tilde{X}$ be regular Lagrangian flows of b and $\tilde{b}$ respectively and denote by L and $\tilde{L}$ their compressibility constants. Then, for every $t\in[0,T]$, we have

$$\|X(t,\cdot)-\tilde{X}(t,\cdot)\|_{L^p(B_r(0))}\le C_t\left|\log(\|b-\tilde{b}\|_{L^1(0,t;L^p(B_R(0)))})\right|^{-1},$$

where $R=r+T\|\tilde{b}\|_\infty$ and the constant C_t only depends on $t,r,\|b\|_\infty,\|\tilde{b}\|_\infty,L,\tilde{L}$ and $\|D_x b\|_{L^1(L^p)}$. Moreover, $C_t\to 0$ as $t\to 0$.

Proof. Let $\delta:=\sqrt{\|b-\tilde{b}\|_{L^1(0,t;L^p(B_R(0)))}}$ with $R=r+T\|\tilde{b}\|_\infty$ and let

$$g(t):=\left[\int_{B_r(0)}\log\left(1+\frac{|X(t,x)-\tilde{X}(t,x)|}{\delta}\right)^p dx\right]^{1/p}.$$

For simplicity of notation, we will omit explicitly writing the dependence on t and x when this will not be necessary. So we will often write X and $\tilde{X}$ in place of $X(t,x)$ and $\tilde{X}(t,x)$. Differentiating in time:

$$\begin{aligned}
&\frac{d}{dt}g(t) \qquad\qquad (4.2.4)\\
&=\frac{1}{p}g(t)^{1-p}\int_{B_r(0)}p\log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1}\frac{\delta}{\delta+|X-\tilde{X}|}\frac{1}{\delta}\frac{X-\tilde{X}}{|X-\tilde{X}|}\left(\frac{d}{dt}X-\frac{d}{dt}\tilde{X}\right)dx\\
&\le\frac{g(t)^{1-p}}{\delta}\int_{B_r(0)}\log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1}\frac{1}{1+|X-\tilde{X}|/\delta}|b(t,X)-\tilde{b}(t,\tilde{X})|dx\\
&\le\underbrace{\frac{g(t)^{1-p}}{\delta}\int_{B_r(0)}\log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1}\frac{1}{1+|X-\tilde{X}|/\delta}|b(t,X)-b(t,\tilde{X})|dx}_{1}\\
&+\underbrace{\frac{g(t)^{1-p}}{\delta}\int_{B_r(0)}\log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1}\frac{1}{1+|X-\tilde{X}|/\delta}|b(t,\tilde{X})-\tilde{b}(t,\tilde{X})|dx}_{2}.
\end{aligned}$$

Considering 1 we have, thanks to Lemma 4.1.2:

$$\begin{aligned}|b(t,X(t,x))-b(t,\tilde{X}(t,x))|\le c_d|X(t,x)-\tilde{X}(t,x)|[&M_{\tilde{R}}|Db|(t,X(t,x))\\ &+M_{\tilde{R}}|Db|(t,\tilde{X}(t,x))]\end{aligned} \tag{4.2.5}$$

for almost every $t\in[0,T]$ and (at fixed t) for almost every $x\in\mathbb{T}^d$, with $\tilde{R}=T(\|b\|_\infty+\|\tilde{b}\|_\infty)$.

Remark 4.2.1.
Notice that the compressibility constraint in the definition of regular Lagrangian flow is crucial for (4.2.5) to hold almost everywhere. Indeed, for almost every t we have that $b(t,\cdot)\in[W^{1,p}(\mathbb{T}^d)]^d$. Then (4.2.5) holds only if $X(t,x),\tilde{X}(t,x)\in\mathbb{T}^d\setminus N_t$, with N_t a null set, which is true since

$$\mathscr{L}^d(X(t,\cdot)^{-1}(N_t))\le L\mathscr{L}^d(N_t)=0\quad\textit{and}\quad\mathscr{L}^d(\tilde{X}(t,\cdot)^{-1}(N_t))\le\tilde{L}\mathscr{L}^d(N_t)=0.$$

Using (4.2.5):

$$\begin{aligned}1\le c_d g(t)^{1-p}\int_{B_r(0)}\log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1}\frac{|X-\tilde{X}|/\delta}{1+|X-\tilde{X}|/\delta}[&M_{\tilde{R}}|Db|(t,X(t,x))\\ &+M_{\tilde{R}}|Db|(t,\tilde{X}(t,x))]dx\end{aligned}$$

$$
\begin{aligned}
&\leq c_d g(t)^{1-p} \int_{B_r(0)} \log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1} [M_{\tilde{R}}|Db|(t,X(t,x)) \\
&\qquad\qquad\qquad\qquad + M_{\tilde{R}}|Db|(t,\tilde{X}(t,x))]dx \\
&= \underbrace{c_d g(t)^{1-p} \int_{B_r(0)} \log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1} M_{\tilde{R}}|Db|(t,X(t,x))dx}_{3} \\
&\quad + \underbrace{c_d g(t)^{1-p} \int_{B_r(0)} \log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1} M_{\tilde{R}}|Db|(t,\tilde{X}(t,x))dx}_{4}.
\end{aligned}
$$

For 3, applying Hölder's inequality with exponents p and $p/(p-1)$ yields

$$
3 \leq c_d g(t)^{1-p} g(t)^{p-1} \left(\int_{B_r(0)} [M_{\tilde{R}}|Db|(t,X(t,x))]^p dx\right)^{1/p},
$$

and changing variables setting $y = X(t,x)$ gives

$$
3 \leq c_d L^{1/p} \left(\int_{B_{r+T\|b\|_\infty}(0)} [M_{\tilde{R}} Db(t,y)]^p dy\right)^{1/p} \leq c_d c_{p,n} L^{1/p} \|Db(t,\cdot)\|_{L^p(B_{R'}(0))}
$$

with $R' = r + 3T\max\{\|b\|_\infty, \|\tilde{b}\|_\infty\}$, thanks to Lemma 4.1.1. With this choice of R', it is easy to see that with the same argument we get

$$
4 \leq c_d c_{p,n} \tilde{L}^{1/p} \|Db(t,\cdot)\|_{L^p(B_{R'}(0))}.
$$

As for 2, we have:

$$
2 \leq \frac{g(t)^{1-p}}{\delta} \int_{B_r(0)} \log\left(1+\frac{|X-\tilde{X}|}{\delta}\right)^{p-1} |b(t,\tilde{X}(t,x)) - \tilde{b}(t,\tilde{X}(t,x))|dx.
$$

First using Hölder's inequality and then changing variables as before:

$$
2 \leq \frac{\tilde{L}^{1/p}}{\delta} \|b(t,\cdot) - \tilde{b}(t,\cdot)\|_{L^p(B_R(0))},
$$

where we recall that $R = r + T\|\tilde{b}\|_\infty$. Inequality (4.2.4) now reads:

$$
\frac{d}{dt} g(t) \leq c_d c_{p,n} \left(L^{1/p} + \tilde{L}^{1/p}\right) \|Db(t,\cdot)\|_{L^p(B_{R'}(0))} + \frac{\tilde{L}^{1/p}}{\delta} \|b(t,\cdot) - \tilde{b}(t,\cdot)\|_{L^p(B_R(0))}.
$$

Thanks to the choice of δ, integrating in time from 0 to $t \in [0,T]$ we get:

$$
g(t) \leq C_t \text{ for all } t \in [0,T],
$$

with C_t being a constant depending on $t, r, L, \tilde{L}, \|D_x b\|_{L^1(L^p(B_{R'}(0))}, \|b\|_\infty, \|\tilde{b}\|_\infty$ and p, and it can be easily seen that $C_t \to 0$ as $t \to 0$. This means that, for every $t \in [0,T]$:

$$
\int_{B_r(0)} \log\left(1+\frac{|X(t,x)-\tilde{X}(t,x)|}{\delta}\right)^p dx \leq C_t^p. \tag{4.2.6}
$$

By Chebyshev's inequality, for every $\eta > 0$ there exists a set K such that $|B_r \setminus K| \leq \eta$ and

$$
\log\left(\frac{|X(t,x)-\tilde{X}(t,x)|}{\delta}+1\right)^p \leq \frac{C_t^p}{\eta} \text{ on } K.
$$

Then, on K:

$$|X(t,x) - \tilde{X}(t,x)|^p \leq \delta^p \exp\left(\frac{C_t p}{\eta^{1/p}}\right).$$

Integrating over B_r:

$$\int_{B_r(0)} |X(t,x) - \tilde{X}(t,x)|^p dx \leq \eta(\|X\|_\infty + \|\tilde{X}\|_\infty)^p + \omega_d r^d \delta^p \exp\left(\frac{C_t p}{\eta^{1/p}}\right), \tag{4.2.7}$$

with $\omega_d := |B_1(0)|$. Choosing $\eta = 2^p C_t^p |\log\delta|^{-p}$ we get

$$\exp\left(\frac{C_t p}{\eta^{1/p}}\right) = \exp\left(\frac{p}{2|\log(\delta)|^{-1}}\right) = \exp\left(-\frac{p}{2}\log(\delta)\right) = \exp(\log(\delta^{-p/2})) = \delta^{-p/2},$$

so (4.2.7) reads as:

$$\int_{B_r(0)} |X(t,x) - \tilde{X}(t,x)|^p dx \leq 2^p C_t^p (\|X\|_\infty + \|\tilde{X}\|_\infty)^p |\log\delta|^{-p} + \omega_d r^d \delta^{p/2}. \tag{4.2.8}$$

Taking the p-th root, and assuming that $\delta = \sqrt{\|b - \tilde{b}\|_{L^1(0,t;L^p(B_R(0)))}}$ is small enough so that $|\log(\delta)|^{-1} \gg \sqrt{\delta}$, we conclude. $\square$

Remark 4.2.2. ***The choice of*** δ
In Proposition 4.2.2 we assumed $\delta := \sqrt{\|b - \tilde{b}\|_{L^1(0,t;L^p(B_R(0)))}}$ *in order to bound* $g(t)$ *with a constant* $C_t \to 0$ *as* $t \to 0$. *If instead we had chosen* $\delta = \|b - \tilde{b}\|_{L^1(0,t;L^p(B_R(0)))}$ *the proof would have still worked, with* $g(t) \leq C$. *In particular, it holds:*

$$\left\|\log\left(1 + \frac{|X(t,\cdot) - \tilde{X}(t,\cdot)|}{\|b - \tilde{b}\|_{L^1(0,t;L^p(B_R(0)))}}\right)\right\|_{L^p(\mathbb{T}^d)} \leq C.$$

As we anticipated in Remark 2.1.5, we are now ready to prove the following.

Proposition 4.2.3. ***Stability of solutions of*** **(TE)*****, revised***
Let $\{b_n\}_{n\in\mathbb{N}}$ *be a sequence of smooth bounded vector fields such that*

$$b_n \to b \text{ in } [L^1((0,T)\times\mathbb{T}^d)]^d \text{ with } b \in [L^1(0,T;W^{1,p}\cap L^\infty(\mathbb{T}^d))]^d$$

and

$$\mathrm{div}_x(b_n) \rightharpoonup \mathrm{div}_x(b) \text{ in } L^1(0,T;L^\infty(\mathbb{T}^d)),$$

with $p > 1$. *Let* $\{\rho_0^n\}_{n\in\mathbb{N}}$ *be equibounded functions converging in* $L^1(\mathbb{T}^d)$ *to* ρ_0. *Then the solutions* ρ_n *of* (TE) *with vector field* b_n *and initial datum* ρ_0^n *converge strongly in* $L^1((0,T)\times\mathbb{T}^d)$ *to the solution* ρ *of* (TE) *with vector field* b *and initial datum* ρ_0.

Proof. We claim that the thesis follows if we prove that ρ_n is a Cauchy sequence in $L^\infty(0,T;L^1(\mathbb{T}^d))$. Indeed, the limit function ρ would be the (unique) solution of (TE) since we could pass to the limit in (2.1.2). That is, for every $\varphi \in C_c^\infty(I\times\mathbb{T}^d)$, if we passed to the limit for $n \to +\infty$ in

$$\int_0^T \int_{\mathbb{T}^d} \rho_n(\partial_t\varphi + \mathrm{div}_x(b_n)\varphi + b_n\cdot\nabla_x\varphi)dxdt + \int_{\mathbb{T}^d}\varphi(0,x)\rho_0^n(x)dx = 0$$

we would have that ρ solves

$$\int_0^T \int_{\mathbb{T}^d} \rho(\partial_t\varphi + \mathrm{div}_x(b)\varphi + b\cdot\nabla_x\varphi)dxdt + \int_{\mathbb{T}^d}\varphi(0,x)\rho_0(x)dx = 0.$$

Notice that we could pass to the limit precisely because we have assumed that ρ_n is a Cauchy sequence with respect to the strong topology in $L^\infty(0,T;L^1(\mathbb{T}^d))$. We thus need to prove that for every $\varepsilon > 0$ there exists N such that

$$\sup_{t\in[0,T]} \|\rho_n(t,\cdot) - \rho_m(t,\cdot)\|_{L^1(\mathbb{T}^d)} \le \varepsilon \quad \forall n,m \ge N.$$

Using the representation formula (1.2.3), we have that

$$\rho_n(t,\cdot) = \rho_0(X_n^{-1}(t,\cdot)),$$

with $X_n^{-1}(t,\cdot)$ the inverse at time t of the (classical) flow $X_n(t,\cdot)$ of b_n. Since weak convergent sequences are bounded, there is a constant $C > 0$ such that $\|\operatorname{div}_x(b)\|_{L^1(0,T;L^\infty(\mathbb{T}^d))} \le C$ for every n, i.e. the compressibility of the flows X_n (and of their inverses X_n^{-1}) is uniformly bounded [5] by C. We now consider a Lipschitz function ρ_0^L such that

$$\|\rho_0^L - \rho_0\|_{L^1(\mathbb{T}^d)} < \frac{\varepsilon}{4C}.$$

Then, for every $t \in [0,T]$ and for every n, m:

$$\begin{aligned}
&\|\rho_n(t,\cdot) - \rho_m(t,\cdot)\|_{L^1(\mathbb{T}^d)} \le \|\rho_0^n(X_n^{-1}(t,\cdot)) - \rho_0^L(X_n^{-1}(t,\cdot))\|_{L^1(\mathbb{T}^d)} \\
&+ \|\rho_0^L(X_n^{-1}(t,\cdot)) - \rho_0^L(X_m^{-1}(t,\cdot))\|_{L^1(\mathbb{T}^d)} + \|\rho_0^L(X_m^{-1}(t,\cdot)) - \rho_0^m(X_m^{-1}(t,\cdot))\|_{L^1(\mathbb{T}^d)} \\
&\le C\|\rho_0^n - \rho_0^L\|_{L^1(\mathbb{T}^d)} + \|\rho_0^L(X_n^{-1}(t,\cdot)) - \rho_0^L(X_m^{-1}(t,\cdot))\|_{L^1(\mathbb{T}^d)} + C\|\rho_0^m - \rho_0^L\|_{L^1(\mathbb{T}^d)},
\end{aligned}$$

where in the first and third addenda we changed variables and used the compressibility bound. Since ρ_0^L is Lipschitz, we have:

$$\|\rho_0^L(X_n^{-1}(t,\cdot)) - \rho_0^L(X_m^{-1}(t,\cdot))\|_{L^1(\mathbb{T}^d)} \le L\|X_n^{-1}(t,\cdot) - X_m^{-1}(t,\cdot)\|_{L^1(\mathbb{T}^d)} \lesssim L|\log(\|b_n - b_m\|_{L^1((0,T)\times\mathbb{T}^d)})|^{-1}$$

by Proposition 4.2.2. It is now trivial to conclude, we just need to choose n, m large enough. □

Remark 4.2.3.
The assumption $p > 1$ is not really necessary and we assumed it for simplicity, so that we could use the quantitative estimate on regular Lagrangian flows of Proposition 4.2.2. In fact, a slight generalisation of the stability result of Theorem 2.2.2 gives (qualitative) stability also if $p = 1$.

4.3 Quantitative estimates for solutions of the continuity equation

We now turn our attention to the PDE counterpart. Namely, we want to give an estimate of how much different two solutions of (CE) are if they have different initial data and they are advected by different vector fields. Just like the Lagrangian estimates in the previous section, the following result will be a sort of blueprint for the estimates in the next chapters, so we will give a proof. For technical reasons [6], and in order to give a result more similar to the one in [144], we consider the dual norm $\|\cdot\|_{W^{-1,1}(\mathbb{T}^d)}$. We also give a less sharp result than [144], where a non-homogeneous continuity equation is considered, but with a proof which is more similar to what we shall use next and which is much simpler. Before proving the stability result let us recall the definition of dual norm of $W^{-1,1}(\mathbb{T}^d)$.

Definition 4.3.1. *Dual norm of* $W^{-1,1}(\mathbb{T}^d)$
Let $\mathcal{L}$ be a continuous linear functional on $W^{1,\infty}(\mathbb{T}^d)$. The dual norm of $\mathcal{L}$ is denoted as $\|\mathcal{L}\|_{W^{-1,1}(\mathbb{T}^d)}$ and it is defined as

[5] Notice that the compressiblity of the inverse of a regular Lagrangian flow is controlled by the $L^1_t L^\infty_x$ norm of the positive part of the divergence.

[6] For a general stability estimate which also accounts for initial data of different mass (i.e. different L^1 norm in the case of positive functions) we cannot use the Wasserstein distance (see Definition 5.3.1) as we will do in the next chapters.

$$\|\mathcal{L}\|_{W^{-1,1}(\mathbb{T}^d)} = \sup\{\mathcal{L}(f) \text{ such that } \|f\|_{W^{1,\infty}(\mathbb{T}^d)} \leq 1\}.$$

As customary we implicitly assume that each function $g \in L^1(\mathbb{T}^d)$ acts as an operator on $W^{1,\infty}(\mathbb{T}^d)$ by integration, namely

$$\|g\|_{W^{-1,1}(\mathbb{T}^d)} = \sup\left\{\int_{\mathbb{T}^d} f(x)g(x)dx \text{ such that } \|f\|_{W^{1,\infty}(\mathbb{T}^d)} \leq 1\right\}.$$

We have the following stability estimate:

Theorem 4.3.1. ***Stability estimate for solutions of* (CE) *(weak), [144]***
Let u_1, u_2 be the solutions of (CE) *with initial data $u_0^1, u_0^2 \in L^q(\mathbb{T}^d)$ with $q \geq 1$ and vector fields $b_1, b_2 \in [L^1(0,T; W^{1,p}(\mathbb{T}^d)) \cap L^\infty((0,T)\times\mathbb{T}^d))]^d$ such that $\mathrm{div}_x(b_1)^-, \mathrm{div}_x(b_2)^- \in L^1(0,T; L^\infty(\mathbb{T}^d))$, with $p > 1$ and $1/p + 1/q \leq 1$. Then it holds:*

$$\|u_1(t,\cdot) - u_2(t,\cdot)\|_{W^{-1,1}(\mathbb{T}^d)} \leq \|u_0^1 - u_0^2\|_{L^1(\mathbb{T}^d)} + C_t \min\{\|u_0^1\|_{L^q(\mathbb{T}^d)}, \|u_0^2\|_{L^q(\mathbb{T}^d)}\} |\log(\|b_1 - b_2\|_{L^1(0,T;L^p(\mathbb{T}^d))})|^{-1},$$

with C_t the constant of Proposition 4.2.2.

Remark 4.3.1.
The same result applies if $u_0^1, u_0^2 \in L^1(\mathbb{T}^d)$, but in that case u_1, u_2 must be assumed to be the Lagrangian solutions of (CE) *(since in general there is no uniqueness of weak solutions outside of the DiPerna-Lions range).*

Proof. By the assumptions on b_1, u_0^1, b_2, u_0^2, the solutions u_1, u_2 are unique, belong to $L_t^\infty L_x^q$ and since the regular Lagrangian flows of b_1, b_2 exist and are unique by the assumptions on the vector fields, the solutions are Lagrangian. That is, they can be represented as the push-forward via the (unique) regular Lagrangian flow of the vector field of the initial datum. So, denoting with X_1 and X_2 the regular Lagrangian flows of b_1 and b_2 respectively, we have:

$$u_1(t,\cdot) = X_1(t,\cdot)_\# u_0^1 \quad \text{and} \quad u_2(t,\cdot) = X_2(t,\cdot)_\# u_0^2 \quad \text{for every } t \in [0,T).$$

By the triangle inequality, for every t it holds:

$$\begin{aligned}
&\|u_1(t,\cdot) - u_2(t,\cdot)\|_{W^{-1,1}(\mathbb{T}^d)} = \|X_1(t,\cdot)_\# u_0^1 - X_2(t,\cdot)_\# u_0^2\|_{W^{-1,1}(\mathbb{T}^d)} \\
&\leq \|X_1(t,\cdot)_\# u_0^1 - X_1(t,\cdot)_\# u_0^2\|_{W^{-1,1}(\mathbb{T}^d)} + \|X_1(t,\cdot)_\# u_0^2 - X_2(t,\cdot)_\# u_0^2\|_{W^{-1,1}(\mathbb{T}^d)}.
\end{aligned}$$

By the push-forward formula, for the first addendum we have:

$$\begin{aligned}
&\|X_1(t,\cdot)_\# u_0^1 - X_1(t,\cdot)_\# u_0^2\|_{W^{-1,1}(\mathbb{T}^d)} = \sup_{\|f\|_{W^{1,\infty}} \leq 1} \int_{\mathbb{T}^d} f(x) dX_1(t,\cdot)_\#(u_0^1 - u_0^2) \\
&= \sup_{\|f\|_{W^{1,\infty}} \leq 1} \int_{\mathbb{T}^d} f(X_1(t,\cdot))(u_0^1(x) - u_0^2(x))dx \leq \|u_0^1 - u_0^2\|_{L^1(\mathbb{T}^d)}.
\end{aligned}$$

For the second addendum:

$$\begin{aligned}
&\|X_1(t,\cdot)_\# u_0^2 - X_2(t,\cdot)_\# u_0^2\|_{W^{-1,1}(\mathbb{T}^d)} = \sup_{\|f\|_{W^{1,\infty}} \leq 1} \int_{\mathbb{T}^d} f(x) dX_1(t,\cdot)_\# u_0^2 - \int_{\mathbb{T}^d} f(x) dX_2(t,\cdot)_\# u_0^2 \\
&= \sup_{\|f\|_{W^{1,\infty}} \leq 1} \int_{\mathbb{T}^d} [f(X_1(t,x)) - f(X_2(t,x))] u_0^2(x) dx \leq \int_{\mathbb{T}^d} |X_1(t,x)) - X_2(t,x)| \, |u_0^2(x)| dx \\
&\leq \|u_0^2\|_{L^q(\mathbb{T}^d)} \|X_1(t,\cdot) - X_2(t,\cdot)\|_{L^p(\mathbb{T}^d)}.
\end{aligned}$$

We just need to apply Proposition 4.2.2 and notice that the argument is symmetric in u_0^1, u_0^2 to have the thesis. □

We now give a result which currently is not present in the literature, but which is easily obtained by the same techniques we already saw: a quantitative estimate for the strong L^p norm of solutions of (CE). Although the proof follows essentially by proving a result very similar to Proposition 4.2.2 (using the techniques of [58]) and we require additional regularity on $\operatorname{div}_x(b)$, we believe that such a result might still be of interest since:

1. The only quantitative estimates in the literature are obtained with respect to very weak topologies (e.g. $W^{-1,1}$ or Wasserstein distance);
2. Incompressible vector fields satisfy the assumptions.

For the sake of simplicity we only consider bounded solutions of (CE) and we provide a L^1 estimate, but L^p generalisations are possible. The key estimate we shall need is the following [7]:

Proposition 4.3.1. ***L^1 stability estimate for $\det(\nabla_{\mathbf{x}}\mathbf{X}(\mathbf{t},\mathbf{x}))$***
Let b and $\tilde b$ be bounded vector fields belonging to $[L^1(0,T;W^{1,p}(\mathbb{T}^d))]^d$ for some $p>1$. Let X and $\tilde X$ be regular Lagrangian flows of b and $\tilde b$ with compressibility $L,\tilde L$ respectively and assume that $\operatorname{div}_x(\tilde b)\in L^1(0,T;L^\infty\cap W^{1,q}(\mathbb{T}^d))$ for some $q>1$ such that $1/p+1/q\le 1$. Let $J(t,x):=\det(\nabla_x X(t,x))$ and $\tilde J(t,x):=\det(\nabla_x\tilde X(t,x))$. Then:

$$\sup_{t\in[0,T]}\|J(t,\cdot)-\tilde J(t,\cdot)\|_{L^1(\mathbb{T}^d)}\le\frac{C}{\left|\log\left[|\log(\|b-\tilde b\|_{L^1(0,T;L^p(\mathbb{T}^d))})|^{-1}+\|\operatorname{div}_x(b)-\operatorname{div}_x(\tilde b)\|_{L^1((0,T)\times\mathbb{T}^d)}\right]\right|},$$

where the constant C depends on the quantities the constant in Proposition 4.2.2 depends on and additionally on $\|J\|_\infty$, $\|\tilde J\|_\infty$, $\|\operatorname{div}_x(\tilde b)\|_{L^1((0,T)\times\mathbb{T}^d)}$ and $\|D_x\operatorname{div}_x(\tilde b)\|_{L^q((0,T)\times\mathbb{T}^d)}$.

Proof. Let $\delta:=\sup_{t\in[0,T]}\|X-\tilde X\|_{L^p(\mathbb{T}^d)}+\|\operatorname{div}_x(b)-\operatorname{div}_x(\tilde b)\|_{L^1((0,T)\times\mathbb{T}^d)}$ and consider the function

$$g(t):=\int_{\mathbb{T}^d}\log\left(1+\frac{|J(t,x)-\tilde J(t,x)|}{\delta}\right)dx.$$

We know that

$$J(t,x)=\exp\left(\int_0^t\operatorname{div}_x(b)(s,X(s,x))ds\right)\text{ and }\tilde J(t,x)=\exp\left(\int_0^t\operatorname{div}_x(\tilde b)(s,\tilde X(s,x))ds\right),$$

so

$$\begin{aligned}
g'(t)&\le\int_{B_r}\frac{|J(t,x)\operatorname{div}_x(b)(t,X(t,x))-\tilde J(t,x)\operatorname{div}_x(\tilde b)(t,\tilde X(t,x))|}{\delta+|J(t,x)-\tilde J(t,x)|}dx\\
&\le\int_{\mathbb{T}^d}\frac{|J(t,x)|\,|\operatorname{div}_x(b)-\operatorname{div}_x(\tilde b)|\circ(t,X(t,x))}{\delta+|J(t,x)-\tilde J(t,x)|}dx+\int_{\mathbb{T}^d}\frac{|J(t,x)-\tilde J(t,x)|\,|\operatorname{div}_x(\tilde b)(t,X(t,x))|}{\delta+|J(t,x)-\tilde J(t,x)|}dx\\
&+\int_{\mathbb{T}^d}\frac{|\tilde J(t,x)|\,|\operatorname{div}_x(\tilde b)(t,X(t,x))-\operatorname{div}_x(\tilde b)(t,\tilde X(t,x))|}{\delta+|J(t,x)-\tilde J(t,x)|}dx.
\end{aligned}$$

The three integrals are bounded. For the first, we use that $J(t,x)$ is uniformly bounded thanks to the condition on $\operatorname{div}_x(b)$, then changing variables (recall that the compressibility is bounded):

$$\int_{\mathbb{T}^d}\frac{|J(t,x)|\,|\operatorname{div}_x(b)-\operatorname{div}_x(\tilde b)|\circ(t,X(t,x))}{\delta+|J(t,x)-\tilde J(t,x)|}dx\lesssim\|J\|_\infty\frac{\|\operatorname{div}_x(b)(t,\cdot)-\operatorname{div}_x(\tilde b)(t,\cdot)\|_{L^1(\mathbb{T}^d)}}{\delta}.$$

The second integral is easier: just use that $|J-\tilde J|/(\delta+|J-\tilde J|)<1$ and change variables to get

$$\int_{\mathbb{T}^d}\frac{|J(t,x)-\tilde J(t,x)|\,|\operatorname{div}_x(\tilde b)(t,X(t,x))|}{\delta+|J(t,x)-\tilde J(t,x)|}dx\lesssim\|\operatorname{div}_x(\tilde b)(t,\cdot)\|_{L^1(\mathbb{T}^d)}.$$

[7]We give a uniform (in time) L^1 estimate in the spirit of [58]. It is clear that with due modifications, like in Proposition 4.2.2, we can localise the estimate in balls, get a L^p estimate and a time dependent constant C_t on the right-hand side.

For the third one, let us denote $f(t,x) = \mathrm{div}_x(\tilde{b})(t,x)$. We can use Lemma 4.1.2 and Lemma 4.1.1 to have:

$$
\begin{aligned}
&\int_{\mathbb{T}^d} \frac{|\tilde{J}(t,x)|\,|\mathrm{div}_x(\tilde{b})(t,X(t,x)) - \mathrm{div}_x(\tilde{b})(t,\tilde{X}(t,x))|}{\delta + |J(t,x) - \tilde{J}(t,x)|} dx \\
&\lesssim \frac{\|\tilde{J}\|_\infty}{\delta} \int_{\mathbb{T}^d} |X(t,x) - \tilde{X}(t,x)| [MD_x f(t,X(t,x)) + MD_x f(t,\tilde{X}(t,x))] dx \\
&\lesssim \frac{\|\tilde{J}\|_\infty}{\delta} \|X(t,\cdot) - \tilde{X}(t,\cdot)\|_{L^p(\mathbb{T}^d)} \|D_x \mathrm{div}_x(\tilde{b})(t,\cdot)\|_{L^q(\mathbb{T}^d)} \le \|\tilde{J}\|_\infty \|D_x \mathrm{div}_x(\tilde{b})(t,\cdot)\|_{L^q(\mathbb{T}^d)}.
\end{aligned}
$$

Integrating in time we find that $g(t) \le C$ for some C. Following the proof of Proposition 4.2.2 and using its thesis, we then find

$$
\begin{aligned}
&\sup_{t\in[0,T]} \|J(t,\cdot) - \tilde{J}(t,\cdot)\|_{L^1(\mathbb{T}^d)} \lesssim C|\log(\delta)|^{-1} \\
&\approx \frac{C}{\left|\log\left[|\log(\|b-\tilde{b}\|_{L^1(0,T;L^p(\mathbb{T}^d))})|^{-1} + \|\mathrm{div}_x(b) - \mathrm{div}_x(\tilde{b})\|_{L^1([0,T]\times\mathbb{T}^d)}\right]\right|}.
\end{aligned}
$$

□

We can now state the stability result:

Proposition 4.3.2. ***Stability estimate for solutions of*** **(CE)** ***(strong)***
Let u_1, u_2 be the solutions of (CE) *with a bounded Lipschitz initial datum u^0 and vector fields $b_1, b_2 \in [L^1(0,T;W^{1,p}(\mathbb{T}^d)) \cap L^\infty((0,T)\times\mathbb{T}^d))]^d$ such that $\mathrm{div}_x(b_1) \in L^1(0,T;L^\infty \cap W^{1,q}(\mathbb{T}^d))$ and $\mathrm{div}_x(b_2) \in L^1(0,T;L^\infty(\mathbb{T}^d))$ with $p > 1$ and $1/p + 1/q \le 1$. Then it holds:*

$$
\begin{aligned}
&\sup_{0\le t<T} \|u_1(t,\cdot) - u_2(t,\cdot)\|_{L^1(\mathbb{T}^d)} \\
&\lesssim \log(\|b_1 - b_2\|_{L^1(0,T;L^p(\mathbb{T}^d))})^{-1} + \frac{1}{\left|\log\left[|\log(\|b-\tilde{b}\|_{L^1(0,T;L^p(\mathbb{T}^d))})|^{-1} + \|\mathrm{div}_x(b) - \mathrm{div}_x(\tilde{b})\|_{L^1([0,T]\times\mathbb{T}^d)}\right]\right|}.
\end{aligned}
$$

Proof. Using the push forward formula and the Lagrangian representation of u_1, u_2 we have:

$$
u_1(t,\cdot) = \frac{u_0}{\det(\nabla_x X_1(t,\cdot))} \circ (X_1^{-1}(t,\cdot)) \text{ and } u_2(t,\cdot) = \frac{u_0}{\det(\nabla_x X_2(t,\cdot))} \circ (X_2^{-1}(t,\cdot)).
$$

We also denote:

$$
J_1(t,x) := \det(\nabla_x X_1(t,x)) \text{ and } J_2(t,x) = \det(\nabla_x X_2(t,x)).
$$

We have:

$$
\begin{aligned}
|u_1(t,x) - u_2(t,x)| &= \left| \frac{u_0(X_1^{-1}(t,x)) J_2(t,X_2^{-1}(t,x)) - u_0(X_2^{-1}(t,x)) J_1(t,X_1^{-1}(t,x))}{J_1(t,X_1^{-1}(t,x)) J_2(t,X_2^{-1}(t,x))} \right| \\
&\le \left| \frac{u_0(X_1^{-1}(t,x)) - u_0(X_2^{-1}(t,x))}{J_1(t,X_1^{-1}(t,x)))} \right| + \|u_0\|_\infty \left| \frac{J_2(t,X_2^{-1}(t,x)) - J_1(t,X_2^{-1}(t,x))}{J_1(t,X_1^{-1}(t,x)) J_2(t,X_2^{-1}(t,x))} \right| \\
&+ \|u_0\|_\infty \left| \frac{J_1(t,X_1^{-1}(t,x)) - J_1(t,X_2^{-1}(t,x))}{J_1(t,X_1^{-1}(t,x)) J_2(t,X_2^{-1}(t,x))} \right|.
\end{aligned}
$$

We now use that the determinants are bounded from below thanks to the assumptions on $\mathrm{div}_x(b_1)$, $\mathrm{div}_x(b_2)$, and that the inverses of the regular Lagrangian flows satisfy the same stability inequalities (namely Proposition 4.2.2 and Proposition 4.3.1) and compressibility estimates since they can be considered as backward flows (see Definition 1.1.2). By the convexity of the exponential and since $(e^x)' = e^x$ we have that

$$|e^x - e^y| \leq \max\{e^x, e^y\}|x - y|,$$

so using the notation $f(t,x) = MD_x \operatorname{div}_x(b_1)(t,x)$ we have by Lemma 4.1.2:

$$\begin{aligned}
&|J_1(t, X_1^{-1}(t,x)) - J_1(t, X_2^{-1}(t,x))| \leq \|J_1\|_\infty \int_0^t |\operatorname{div}_x(b_1)(s, X_1^{-1}(t-s,x)) - \operatorname{div}_x(b_1)(s, X_2^{-1}(t-s,x))|ds \\
&\leq \|J_1\|_\infty \underbrace{\left(\int_0^t |f(s, X_1^{-1}(t-s,x)) + f(s, X_2^{-1}(t-s,x))|\,|X_1^{-1}(t-s,x) - X_2^{-1}(t-s,x)|ds\right)}_{:=\eta(t,x)} \\
&= \|J_1\|_\infty \eta(t,x).
\end{aligned}$$

Thus:

$$\begin{aligned}
&|u_1(t,x) - u_2(t,x)| \lesssim \|J_1^{-1}\|_\infty |X_1^{-1}(t,x) - X_2^{-1}(t,x)| + \|u_0\|_\infty \|(J_1J_2)^{-1}\|_\infty |J_2(t, X_2^{-1}(t,x)) - J_1(t, X_2^{-1}(t,x)| \\
&+ \|u_0\|_\infty \|J_1\|_\infty \|(J_1J_2)^{-1}\|_\infty \eta(t,x),
\end{aligned}$$

with C depending on the same quantities of the constants in Proposition 4.2.2 and Proposition 4.3.1, and moreover depending also on $\|u_0\|_{L^\infty}$ and $\|(J_1J_2)^{-1}\|_\infty$. Integrating over $\mathbb{T}^d$ and using the bounded compressibility of the backward flow and Hölder's inequality for estimating $\int_{\mathbb{T}^d} \eta(t,x)dx$:

$$\sup_{t\in[0,T]} \|u_1(t,\cdot) - u_2(t,\cdot)\|_{L^1(\mathbb{T}^d)} \lesssim \sup_{t\in[0,T]} \|X_1^{-1}(t,\cdot) - X_2^{-1}(t,\cdot)\|_{L^1(\mathbb{T}^d)} + \sup_{t\in[0,T]} \|J_1(t,\cdot) - J_2(t,\cdot)\|_{L^1(\mathbb{T}^d)}.$$

We conclude by Proposition 4.2.2 and Proposition 4.3.1.

□

Remark 4.3.2.
It is not difficult to see that we may obtain an equivalent result for a non-Lipschitz unbounded initial datum and for different initial data u_0^1, u_0^2. In the former case, we would have an additional term $\|u_0 - u_0^L\|_{L^1(\mathbb{T}^d)}$ where u_0^L is a bounded Lipschitz function approximating u_0. In the latter, we would have an additional error term $\|u_0^1 - u_0^2\|_{L^1(\mathbb{T}^d)}$.

Chapter 5

Numerical approximation of regular Lagrangian flows

In this chapter we address the problem of approximating both the regular Lagrangian flow and the solution of the continuity equation with respect to a non-regular vector field b. In order to make this chapter more self-contained we recall here the continuity equation: the initial value problem we consider is

$$\begin{cases} \partial_t u + \operatorname{div}_x(bu) = 0 & \text{in } (0,T)\times\mathbb{T}^d \\ u(0,\cdot) = u_0 & \text{in } \mathbb{T}^d \end{cases}. \tag{PDE}$$

In recent years the topic of approximating the solution of (PDE) in presence of non-Lipschitz velocity fields with numerical methods has started to be investigated. To the best of our knowledge, the first works in this direction have been [159] and [31], where using a discontinuous Galerkin method and a finite volume scheme respectively they prove strong convergence of approximate solutions, although without quantitative rates. It is worth mentioning that in [31] a rate of $1/2$, based on numerical simulations, is conjectured, which matches the rate for smoother vector fields [1]. The rate was proved to be optimal, although only for Lipschitz velocity fields, in [123, 124]. The first quantitative estimate of the error in presence of a vector field with Sobolev regularity (although only in a weak sense, i.e. in logarithmic Wasserstein distance, introduced in this context in [144]) was obtained in [142], where an *explicit* finite volume upwind scheme is considered, which requires (for technical reasons) the use of a Cartesian mesh [2] to discretise the spatial domain. Such results have been extended to more general meshes [143] by using an *implicit* method, and recently they have been adapted to include diffusion [128, 129]. Recent developments concerning the numerical approximation of (PDE) have also been obtained in [19, 96]. In particular, in [19] the strong convergence (i.e. convergence in the L^p norm) of approximate solutions to non-linear continuity equations of the form

$$\partial_t u + \operatorname{div}_x(uf(u)) = 0$$

is proved for various approximation schemes. For instance, strong convergence of the Lax–Friedrichs scheme [108, 113, 114] and of the upwind scheme is established. The argument relies on bounding the semi-norms of the approximate solutions on a precise logarithmic Sobolev space [18] which embeds compactly in L^p. We refer the interested reader to [19, Theorem 8] pointing out that, while not explicitly derived, the results obtained in [19] allow to get a quantitative rate of convergence (see [19, Corollary 11] and the discussion after that). However, as in [142], the main restriction is that the estimates are only valid in a Cartesian mesh. The issue has been partially solved in [96], where the authors obtain similar results for the continuity equation on non-Cartesian meshes with some additional hypotheses (they require some periodicity condition). Moreover, in [96] the authors manage to treat a system of equations where the density u solving (PDE) and the velocity field b are coupled in a non linear way [96, Theorem 1.7]. We stress that, since the continuity equation is often part of many different physical models, studying the coupling of (PDE) with other equations is crucial in many different problems from physics to biology. A non-exhaustive list comprehends [81, 82] for compressible Navier-Stokes, [16, 74, 80] for compressible Stokes, [153] for swarming and [42] for chemotaxis, see also [96] and the references therein.

[1]The method is actually of first order, but the rate of convergence falls to 1/2 if the initial datum is not regular [104, 67].

[2]I.e. a mesh whose elements are isometric axis-parallel rectangular boxes.

To the best of our knowledge, so far every approximation scheme has employed an Eulerian point of view of (PDE). We will instead use a Lagrangian one, managing to prove quantitative estimates, although only in a weak topology, for non-Cartesian meshes without *any* regularity assumptions on the mesh, requiring only a bound on the diameter of the cells. This seems to be the first *explicit* method which gives quantitative estimates on a non-Cartesian mesh (we recall that in [143] they used an *implicit* scheme).

Another significant advantage of our method is the lack of a CFL condition that must be satisfied in order to have stability. As we shall see in the proofs, the errors will be controlled either by the sum of terms which depend on the time step Δt and on the mesh size Δx separately, or by terms depending on $\max\{\Delta t, \Delta x\}$. This seems to be a clear advantage of the Lagrangian point of view, since in the previously cited works, which studied the problem from an Eulerian perspective, a CFL condition was always needed (see Remark 5.4.2). Finally, we note that a parallelisation of our method is straightforward.

5.1 Setting and main assumptions

We will work with functions satisfying the following *main assumptions*:

$$\begin{cases} u_0 \in L^q(\mathbb{T}^d) & \\ b \in [L^1(0,T;W^{1,p}(\mathbb{T}^d)) \cap L^\infty((0,T)\times\mathbb{T}^d)]^d & \text{with } 1/p + 1/q \leq 1,\ p > 1 \\ \operatorname{div}_x(b)^- \in L^\infty((0,T)\times\mathbb{T}^d) & \text{if } 1 < p \leq d \\ \operatorname{div}_x(b)^- \in L^1(0,T;L^\infty(\mathbb{T}^d)) & \text{if } p > d \end{cases} . \tag{MA}$$

Remark 5.1.1.
Our results are stated on the d-dimensional torus $\mathbb{T}^d$ for simplicity. However, with some care, they can be extended to more general ambient spaces.

Notice that when $1 < p \leq d$ we are asking a little more than what is asked in [70, 7] for having well-posedness of (PDE). Namely, we are not asking that $\operatorname{div}_x(b)^- \in L^1(0,T;L^\infty(\mathbb{T}^d))$, but rather $\operatorname{div}_x(b)^- \in L^\infty((0,T)\times\mathbb{T}^d)$. This stronger assumption is technical and it will be crucial in our proof of Lemma 5.2.1. The idea is to use Lagrangian estimates to get estimates at the Eulerian level (i.e. for the solution u of (PDE)). The key observation is that under the assumptions (MA) the solution u of (PDE) is Lagrangian, i.e. it can be represented as

$$u(t,\cdot) = \Phi(t,\cdot)_\# u_0,$$

with u_0 the initial datum and Φ the regular Lagrangian flow associated to the velocity b, as defined in Definition 2.2.4. The idea is to approximate the solution u with

$$u_E(t,\cdot) = \Phi_E(t,\cdot)_\# \mu,$$

where Φ_E is an approximation of Φ and μ is an approximation of u_0. This method is in principle quite flexible, in the sense that any "good" approximation of Φ and u_0 should yield a "good" approximation of u in a sense which we shall specify more precisely later. In particular we approximate u in two ways, using different approximations for both Φ and u_0. In the first one we consider a mesh of the domain with some mild assumptions on the sets $\{Q_i\}_{i=1}^N$ of the mesh, and for every set Q_i we choose (according to some probability distribution) a point $x_i \in Q_i$ and a weight $M_i \geq 0$ so that $\mu = \sum_{i=1}^M M_i \delta_{x_i}$ approximates the initial datum u_0. As for the approximation of the regular Lagrangian flow Φ we construct, for every point $x \in \mathbb{T}^d$, a piecewise affine map $\Phi_E(\cdot, x)$ obtained by a (sort of) explicit Euler method applied to the velocity field b. The precise definition of Φ_E is given in Definition 5.2.1. The approximation u_E is therefore a weighted sum of Dirac deltas whose positions evolve in time in a piecewise linear fashion. We prove that u_E approximates in Wasserstein distance the solution u of (PDE), although only in probability (see Theorem 5.3.2, Theorem 5.3.3 and Proposition 5.3.1 for the precise statements). We do not know if, at least in some cases, similar estimates hold for *every* choice of the initial positions of the Dirac deltas, i.e. in a deterministic sense (see Remark 5.3.4). We also consider a different approximation $\bar{\Phi}_E$ of the flow (see Definition 5.4.1) which, if $p > d$ and considering an approximation of the initial datum in the strong

L^1 norm, provides a deterministic approximation of u in Wasserstein distance (see Theorem 5.4.1). The novelty is that this an *explicit* method which can be performed on unstructured meshes, it does not need to satisfy a CFL condition (see Remark 5.4.2), it provides an easily parallelisable algorithm and it uses the Lagrangian formulation of the problem, whereas in previous works the Eulerian point of view was always employed.

Let us now discuss briefly the organisation of this chapter. Section 5.2 is dedicated to proving Theorem 5.2.1, where we give a novel $L_t^\infty(L_x^p)$ stability estimate for the difference of the regular Lagrangian flow Φ of a velocity field b satisfying (MA) and an approximation Φ_E of such flow, in the spirit of the stability estimates between regular Lagrangian flows derived in [58] (see also Proposition 4.2.2). In Section 5.3 we use such a stability estimate to approximate, in Wasserstein distance, the solution u of the continuity equation (PDE) via the push-forward through Φ_E of suitably chosen Dirac deltas. The estimates are proved both in 1-Wasserstein distance and logarithmic Wasserstein distance. In the latter case in particular we show that in expected value the rate of convergence is of order $1/2$ if $1 < p \leq d$, as the rate obtained in [142], and it is better than that if $p > d$ (see Theorem 5.3.2 and Theorem 5.3.3 for the details). We believe that the logarithmic rates we obtain are optimal for the range $1 < p \leq d$ (as suggested also in [142]), but the case $p > d$ does not give any precise rate. In section 5.4 we consider, for technical reasons explained in Remark 5.4.1, only velocity fields $b \in [L^1(0,T; W^{1,p} \cap L^\infty(\mathbb{T}^d)]^d$ such that $p > d$ and we assume to approximate the initial datum strongly in L^1. Under these stronger assumptions and using a slightly different approximation $\bar{\Phi}_E$ of the regular Lagrangian flow, we are able to prove approximation results for the solution of (PDE) (similar to the ones of section 5.3) which hold deterministically. We refer the reader to Theorem 5.4.1 for the details. In section 5.5 we give approximation results for the regular Lagrangian flow and show numerical experiments using a θ-method, i.e. a first-order method which the explicit Euler method is a particular case of. Finally, in section 5.6 we summarise the results obtained and highlight some open questions and future research perspectives.

5.2 Stability estimates

The main result of this section, i.e. Theorem 5.2.1, is a stability estimate in the spirit of Proposition 4.2.2 which, although interesting on its own, can be used as a tool to prove approximation results for the solution of the continuity equation in a way similar to what we saw in Theorem 4.3.1. Let us immediately state the main result of this section, namely a stability estimate between the regular Lagrangian flow Φ and an approximation Φ_E of this flow, which will be introduced shortly after the statement.

Theorem 5.2.1. ***L^p stability estimate between Φ and Φ_E***
Let b satisfy assumptions (MA). *If Φ is the regular Lagrangian flow of b and Φ_E is defined as in Definition 5.2.1 with $\delta = \sqrt{\Delta t}$ if $1 < p \leq d$, we have*

$$\|\Phi(t,\cdot) - \Phi_E(t,\cdot)\|_{L^p(\mathbb{T}^d)} \lesssim C_t |\log(\Delta t)|^{-1},$$

where $C_t \to 0$ as $t \to 0$. Moreover, if $p > d$, the same estimate holds even if $\operatorname{div}_x(b)^- \in L^1(0,T; L^p(\mathbb{T}^d))$.

The vector field Φ_E is explicitly constructed by a (sort of) explicit Euler method, and we now give its precise definition.

Definition 5.2.1. ***Euler approximation of a regular Lagrangian flow***
Let b satisfy conditions (MA) *with $1 < p \leq d$, and let*

$$b_\delta(t,\cdot) := \eta_\delta * b(t,\cdot) \tag{5.2.1}$$

for some $\delta > 0$, where η is a standard mollifier in $\mathbb{R}^d$, $\eta_\delta(x) := \delta^{-d}\eta(x/\delta)$ and the convolution is in space only. We denote by Φ_E (where "E" stands for "Euler") the so-called Euler approximation of the regular Lagrangian flow Φ of b, with time step Δt and regularisation parameter δ. For $t \in (t_n, t_{n+1})$ we define:

$$\Phi_E(t,x) := \Phi_E(t_n,x) + (t - t_n) \fint_{t_n}^{t_{n+1}} b_\delta(s, \Phi_E(t_n,x))ds, \tag{5.2.2}$$

where $0 = t_0 < t_1 < \cdots < t_M = T$ *is a partition of* $[0, T]$ *with* $t_{i+1} - t_i = \Delta t$ *and* $\Phi_E(0, x) = x\ \forall x \in \mathbb{T}^d$. *If* $p > d$, *we define* Φ_E *directly with* b *as*

$$\Phi_E(t, x) := \Phi_E(t_n, x) + (t - t_n) \fint_{t_n}^{t_{n+1}} b(s, \Phi_E(t_n, x))ds. \tag{5.2.3}$$

We remark that, in order to keep the notation as light as possible, we omit explicitly writing the dependence of Φ_E *on* Δt *and (eventually)* δ.

In order to prove Theorem 5.2.1, we need some preliminary estimates. A key result is the L^p stability estimate in Proposition 4.2.2, with which we can easily prove the following:

Corollary 5.2.1. ***Error of regularisation***
Consider Φ *the regular Lagrangian flow of a vector field* b *satisfying* (MA), *and let* Φ_δ *be the (classical) flow of* b_δ *defined in* (5.2.1). *Then, for every* $t \in [0, T]$:

$$\|\Phi(t, \cdot) - \Phi_\delta(t, \cdot)\|_{L^p(\mathbb{T}^d)} \lesssim C_t |\log(\|b - b_\delta\|_{L^1(0,t;L^p(\mathbb{T}^d))})|^{-1} \le C_t |\log(\delta)|^{-1},$$

with $C_t \to 0$ *as* $t \to 0$.

Proof. For the first inequality, use the L^p estimate between regular Lagrangian flows in Proposition 4.2.2. For the second one notice that denoting with B_1 the ball centered in 0 and of unitary radius, using Lemma 4.1.2:

$$\begin{aligned}
&\left|b(t, x) - b_\delta(t, x)\right| = \left|b(t, x) - \int_{\mathbb{T}^d} \eta_\delta(z) b(t, x + z)dz\right| \\
&= \left|b(t, x) - \int_{B_1} \eta(y) b(t, x + \delta y)dy\right| \le \int_{B_1} \eta(y)|b(t, x) - b(t, x + \delta y)|dy \\
&\lesssim \delta \int_{B_1} \eta(y)[M_\delta|Du|(t, x) + M_\delta|Du|(t, x + \delta y)]dy \\
&= \delta M_\delta|Du|(t, x) + \delta[\eta_\delta * M_\delta|Du|](t, x).
\end{aligned}$$

Take the L^p norm on both sides and conclude using Young's convolution inequality, the boundedness of the maximal operator, and finally integrating in time. We thus have

$$\begin{aligned}
&\|b - b_\delta\|_{L^1(0,t;L^p(\mathbb{T}^d))} \lesssim \delta \implies \log(\|b - b_\delta\|_{L^1(0,t;L^p(\mathbb{T}^d))}) \lesssim \log(\delta) \\
&\implies |\log(\|b - b_\delta\|_{L^1(0,t;L^p(\mathbb{T}^d))})| \gtrsim |\log(\delta)| \\
&\implies |\log(\|b - b_\delta\|_{L^1(0,t;L^p(\mathbb{T}^d))})|^{-1} \lesssim |\log(\delta)|^{-1}.
\end{aligned}$$

□

We now need to estimate $\|\Phi_\delta(t, \cdot) - \Phi_E(t, \cdot)\|_{L^p(\mathbb{T}^d)}$ uniformly in t, but since Φ_E is not properly a flow we cannot use the stability estimate in Proposition 4.2.2. We need the following propositions as a key tool to treat the case $1 < p \le d$. As we will see in the proof of Proposition 5.2.3, this will not be needed if $p > d$.

Proposition 5.2.1. ***Compressibility estimate for*** Φ_E
Consider Φ_E *defined as in* (5.2.2) *with* $\delta = \sqrt{\Delta t}$. *Then* $\exists c > 0$ *not dependent on* Δt *such that*

$$\det(\nabla_x \Phi_E(t, x)) > c \text{ for every } (t, x) \in [0, T] \times \mathbb{T}^d.$$

Proof. Define $\Phi_n(t, x)$ in $[0, \Delta t] \times \mathbb{T}^d$ as

$$\Phi_n(t, x) := x + t \fint_{t_n}^{t_{n+1}} b_\delta(s, x)ds.$$

Notice that Φ_E can be seen as a composition of such Φ_n's, i.e. if $t \in [t_n, t_{n+1}]$:

$$\Phi_E(t,x) = \Phi_n(t-t_n, \Phi_{n-1}(\Delta t, \Phi_{n-2}(\dots, \Phi_0(\Delta t, x)))\dots).$$

Differentiating in space:

$$\nabla_x \Phi_E(t,x) = \nabla_x \Phi_n(t-t_n, \dots)\nabla_x \Phi_{n-1}(\Delta t, \dots)\dots \nabla_x \Phi_0(\Delta t, x),$$

and by Binet:

$$\det[\nabla_x \Phi_E(t,x)] = \underbrace{\det[\nabla_x \Phi_n(t-t_n,\dots)]\det[\nabla_x \Phi_{n-1}(\Delta t,\dots)]\dots\det[\nabla_x \Phi_0(\Delta t,x)]}_{\approx 1/\Delta t \text{ times}}.$$

We will uniformly bound $\det[\nabla_x \Phi_0]$ from below, but the argument is the same for every Φ_j (the only difference is the time interval). We have

$$\begin{aligned}\det\left(\nabla_x \Phi_0(\Delta t, x)\right) &= \det\left(I + \Delta t \fint_0^{\Delta t} \nabla_x b_\delta(s,x)ds\right)\\ &= \det(I+\Delta t A), \quad \text{where } A := \fint_0^{\Delta t} \nabla_x b_\delta(s,x)ds.\end{aligned}$$

It is known (see [85, Chapitre I ,section 3 , equations (4) and (6)]) that

$$\det(I+\Delta t A) = 1 + \Delta t[tr(A)] + \sum_{j=2}^d (\Delta t)^j \sigma_j(\lambda_1,\dots,\lambda_d),$$

where σ_j are symmetric polynomials of degree j and λ_i are the eigenvalues of A. By our definition each entry of the matrix A is bounded by $\approx 1/\sqrt{\Delta t}$ since b_δ is $1/\sqrt{\Delta t}$-Lipschitz, and thus $|\lambda_i| \lesssim 1/\sqrt{\Delta t}$ for every i. Then, $|\sigma_j(\lambda_1,\dots,\lambda_d)| \lesssim (\Delta t)^{-j/2}$. Since we also have that $tr(A)^- = \operatorname{div}_x(b_\delta)^- \in L^\infty((0,T)\times\mathbb{T}^d)$, there exists $C>0$ such that:

$$\det(I+\Delta t A) \geq 1 - \Delta t[\operatorname{div}_x(b)^-] - \sum_{j=2}^d (\Delta t)^j(\Delta t)^{-j/2} \geq 1 - C\Delta t. \tag{5.2.4}$$

Multiplying $\approx 1/\Delta t$ times such determinants we can estimate $\det(\nabla_x \Phi_E)$ uniformly from below, and the proof is complete. □

Remark 5.2.1.
In the following we will always assume $\delta = \sqrt{\Delta t}$. In view of the proof of Proposition 5.2.1 it is clear that other choices are possible. However, as we shall see in Theorem 5.3.3, the choice of $\delta = \sqrt{\Delta t}$ will be crucial for obtaining a rate of convergence of order $1/2$ if we consider a logarithmic Wasserstein distance, assuming that the space discretization scale Δx is equal to Δt.

Proposition 5.2.2. ***Injectivity and Lipschitz continuity of*** $\Phi_E(t,\cdot)$
Let Φ_E be defined as in (5.2.2) *and let $\delta = \sqrt{\Delta t}$. Then $\exists C_{\Delta t,T} > 0$ such that*

$$C_{\Delta t,T}^{-1}|x-y| \leq |\Phi_E(t,x) - \Phi_E(t,y)| \leq C_{\Delta t,T}|x-y| \text{ for all } t\in[0,T].$$

In particular, $\Phi_E(t,\cdot)$ is injective and Lipschitz continuous.

Proof. Consider two points $x,y\in\mathbb{T}^d$, with $|x-y| = \ell$. At the first step in the definition of Φ_E (i.e. up to time $t=\Delta t$) we have, since b_δ is $1/\sqrt{\Delta t}$-Lipschitz:

$$|\Phi_E(\Delta t,x) - \Phi_E(\Delta t,y)| \leq |x-y| + \int_0^{\Delta t} |b_\delta(s,x) - b_\delta(s,y)|ds \lesssim \ell(1+\sqrt{\Delta t}).$$

Iterating $T/\Delta t$ times we get a Lipschitz constant $C_{\Delta t,T} = \exp(T/\sqrt{\Delta t})$, and with the same argument we get an estimate from below with $C_{\Delta t,T}^{-1} = \exp(-T/\sqrt{\Delta t})$. Injectivity follows immediately. □

We are now ready to prove the following proposition, whose proof which makes use of the concavity of the logarithm is strongly inspired by the proof of [143, Lemma 6]:

Proposition 5.2.3. ***L^p stability estimate for Φ_δ and Φ_E***
Let Φ_E be defined as in (5.2.2) *with b satisfying assumptions* (MA)*, and let $\delta = \sqrt{\Delta t}$. Then:*

$$\|\Phi_\delta(t,\cdot) - \Phi_E(t,\cdot)\|_{L^p(\mathbb{T}^d)} \lesssim C_t |\log(\Delta t)|^{-1} \text{ for every } t \in [0,T],$$

where $C_t \to 0$ as $t \to 0$ and Φ_δ is the flow of b_δ. If $b \in L^1(0,T; W^{1,p}(\mathbb{T}^d))$ with $p > d$, then

$$\|\Phi(t,\cdot) - \Phi_E(t,\cdot)\|_{L^p(\mathbb{T}^d)} \lesssim C_t |\log(\Delta t)|^{-1} \text{ for every } t \in [0,T],$$

where $C_t \to 0$ as $t \to 0$ and Φ_E is defined as in (5.2.3)*. Moreover, if $p > d$ the conclusion holds even if we only require $\operatorname{div}_x(b)^- \in L^1(0,T; L^\infty(\mathbb{T}^d))$.*

Proof. In the first part of the proof we will work without loss of generality with Φ_δ with the assumption $1 < p \le d$, since the argument is identical if $p > d$. The difference will appear only at the end of the proof, where we differentiate the two cases. We want to prove that

$$\left\| \log\left(1 + \frac{|\Phi_\delta(t,\cdot) - \Phi_E(t,\cdot)|}{\Delta t}\right) \right\|_{L^p(\mathbb{T}^d)} \le C_t, \tag{5.2.5}$$

with a constant C_t not dependent on Δt such that $C_t \to 0$ as $t \to 0$. Once we prove (5.2.5), the conclusion follows as in the last part of the proof of Proposition 4.2.2. Notice that, by the concavity of the logarithm and by the fact that $\log(1+x)$ is an increasing monotone function, for every $a, b \ge 0$ it holds

$$\log(1+b) - \log(1+a) \le \frac{d}{dx}\Big[\log(1+x)\Big]_{x=a} |b-a| = \frac{|b-a|}{1+a}.$$

In particular, let $t = t_n$ for simplicity. We have (since the estimate is pointwise we omit the dependence on x to keep the notation as light as possible):

$$\begin{aligned}
&\log\left(1 + \frac{|\Phi_\delta(t_n) - \Phi_E(t_n)|}{\Delta t}\right) - \log\left(1 + \frac{|\Phi_\delta(t_{n-1}) - \Phi_E(t_{n-1})|}{\Delta t}\right) \\
&\le \frac{\Delta t}{|\Phi_\delta(t_{n-1}) - \Phi_E(t_{n-1})| + \Delta t} \frac{|\,|\Phi_\delta(t_n) - \Phi_E(t_n)| - |\Phi_\delta(t_{n-1}) - \Phi_E(t_{n-1})|\,|}{\Delta t} \\
&\le \frac{|\Phi_\delta(t_n) - \Phi_\delta(t_{n-1}) - (\Phi_E(t_n) - \Phi_E(t_{n-1}))|}{|\Phi_\delta(t_{n-1}) - \Phi_E(t_{n-1})| + \Delta t} \\
&= \frac{|\int_{t_{n-1}}^{t_n} b_\delta(s, \Phi_\delta(s)) - b_\delta(s, \Phi_E(t_{n-1})) ds|}{|\Phi_\delta(t_{n-1}) - \Phi_E(t_{n-1})| + \Delta t}.
\end{aligned}$$

Iterating the estimate above up to $t_0 = 0$ we get by telescopic summing:

$$\begin{aligned}
&\log\left(1 + \frac{|\Phi_\delta(t_n) - \Phi_E(t_n))|}{\Delta t}\right) \le \sum_{j=0}^{n-1} \frac{|\int_{t_j}^{t_{j+1}} b_\delta(s, \Phi_\delta(s)) - b_\delta(s, \Phi_E(t_j)) ds|}{|\Phi_\delta(t_j) - \Phi_E(t_j)| + \Delta t} \qquad (5.2.6)\\
&\le \sum_{j=0}^{n-1} \frac{|\int_{t_j}^{t_{j+1}} b_\delta(s, \Phi_\delta(s)) - b_\delta(s, \Phi_\delta(t_j)) ds| + |\int_{t_j}^{t_{j+1}} b_\delta(s, \Phi_\delta(t_j)) - b_\delta(s, \Phi_E(t_j)) ds|}{|\Phi_\delta(t_j) - \Phi_E(t_j)| + \Delta t} \\
&\le \underbrace{\sum_{j=0}^{n-1} \frac{\int_{t_j}^{t_{j+1}} |b_\delta(s, \Phi_\delta(s)) - b_\delta(s, \Phi_\delta(t_j))| ds}{|\Phi_\delta(t_j) - \Phi_E(t_j)| + \Delta t}}_{=1} + \underbrace{\sum_{j=0}^{n-1} \frac{\int_{t_j}^{t_{j+1}} |b_\delta(s, \Phi_\delta(t_j)) - b_\delta(s, \Phi_E(t_j))| ds}{|\Phi_\delta(t_j) - \Phi_E(t_j)| + \Delta t}}_{=2}.
\end{aligned}$$

Now, notice that by Lemma 4.1.2:

$$
\begin{aligned}
&\int_{t_j}^{t_{j+1}} |b_\delta(s,\Phi_\delta(s)) - b_\delta(s,\Phi_\delta(t_j))|ds \\
&\lesssim \int_{t_j}^{t_{j+1}} |\Phi_\delta(s) - \Phi_\delta(t_j)|\ [M(|Db_\delta|)(s,\Phi_\delta(s)) + M(|Db_\delta|)(s,\Phi_\delta(t_j))]ds \\
&= \int_{t_j}^{t_{j+1}} \left|\int_{t_j}^{s} b_\delta(\tau,\Phi_\delta(\tau)d\tau\right| [M(|Db_\delta|)(s,\Phi_\delta(s)) + M(|Db_\delta|)(s,\Phi_\delta(t_j))]ds \\
&\le \|b_\delta\|_\infty \Delta t \int_{t_j}^{t_{j+1}} [M(|Db_\delta|)(s,\Phi_\delta(s)) + M(|Db_\delta|)(s,\Phi_\delta(t_j))]ds.
\end{aligned}
$$

Then, since $\Delta t/(|\Phi_\delta(t_j) - \Phi_E(t_j)| + \Delta t) \le 1$, we have

$$
\begin{aligned}
1 &\lesssim \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} [M(|Db_\delta|)(s,\Phi_\delta(s)) + M(|Db_\delta|)(s,\Phi_\delta(t_j))]ds \\
&= \int_0^T M(|Du|)(s,\Phi_\delta(s))ds + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} M(|Db_\delta|)(s,\Phi_\delta(t_j))ds.
\end{aligned}
$$

Considering the L^p norm:

$$
\left\| \int_0^T M(|Du|)(s,\Phi_\delta(s))ds + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} M(|Db_\delta|)(s,\Phi_\delta(t_j))ds \right\|_{L^p(\mathbb{T}^d)} \tag{5.2.7}
$$

$$
\begin{aligned}
&\le \int_0^T \|M(|Db_\delta|)(s,\Phi_\delta(s))\|_{L^p(\mathbb{T}^d)} + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|M(|Db_\delta|)(s,\Phi_\delta(t_j))\|_{L^p(\mathbb{T}^d)} ds \\
&\lesssim \|M(Db_\delta)\|_{L^1(0,T;L^p(\mathbb{T}^d))} \lesssim \|Du\|_{L^1(0,T;L^p(\mathbb{T}^d))},
\end{aligned}
$$

where in the last inequalities we used the L^∞ bound on the compressibility to change variables in each addendum, the boundedness of the maximal operator and Young's convolution inequality. As for 2, we need to differentiate the cases $1 < p \le d$ and $p > d$.

- If $1 < p \le d$, we repeat the argument above. We have

$$
\begin{aligned}
&\int_{t_j}^{t_{j+1}} |b_\delta(s,\Phi_\delta(t_j)) - b_\delta(s,\Phi_E(t_j))|ds \\
&\lesssim |\Phi_h(t_j) - \Phi_E(t_j)| \int_{t_j}^{t_{j+1}} [M(|Db_\delta|)(s,\Phi_\delta(t_j)) + M(|Db_\delta|)(s,\Phi_E(t_j))]ds,
\end{aligned}
$$

  and we can estimate $|\Phi_h(t_j) - \Phi_E(t_j)|/(|\Phi_h(t_j) - \Phi_E(t_j)| + \Delta t) \le 1$. Then, after summing over j:

$$
2 \lesssim \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} [M(|Db_\delta|)(s,\Phi_\delta(t_j)) + M(|Db_\delta|)(s,\Phi_E(t_j))]ds.
$$

  After considering the L^p norm the estimate goes as in (5.2.7). For the second addendum, notice that by Proposition 5.2.2 the map Φ_E is Lipschitz and injective (so we are allowed to change variables in the usual way) and by Proposition 5.2.1 the compressibility is controlled, thus the change of variables does not cause any problem as in (5.2.7) when we were dealing with Φ_δ, and we can conclude.

- If $p > d$, we recall that we work directly with Φ, and not with Φ_δ. In this case we use Lemma 4.1.3 with $\tilde{p} \in (d,p)$ to get:

$$\begin{aligned}\frac{\int_{t_j}^{t_{j+1}} |b(s,\Phi(t_j)) - b(s,\Phi_E(t_j))|ds}{|\Phi(t_j) - \Phi_E(t_j)| + \Delta t} &\lesssim \frac{|\Phi(t_j) - \Phi_E(t_j)| \int_{t_j}^{t_{j+1}} f(s,\Phi(t_j))ds|}{|\Phi(t_j) - \Phi_E(t_j)| + \Delta t} \\ &\leq \int_{t_j}^{t_{j+1}} f(s,\Phi(t_j))ds,\end{aligned} \tag{5.2.8}$$

  where

$$f(s,x) := [M(|Du|^{\tilde{p}})(s,x)]^{1/\tilde{p}} \in L^1(0,T;L^{\tilde{p}}(\mathbb{T}^d)). \tag{5.2.9}$$

  To conclude we just need to sum over j, consider the L^p norm and use the compressibility estimate on Φ to change variables. Notice that $\tilde{p} \in (d,p)$ is needed to have $|Db_\delta(s,\cdot)| \in L^{p/\tilde{p}}(\mathbb{T}^d)$ with $p/\tilde{p} > 1$ in order to have boundedness of the maximal operator. Notice also that in this case we do not need any compressibility estimate for Φ_E since the argument of f is $\Phi(t_j) = \Phi(t_j,\cdot)$, whose compressibility is controlled even under the weaker assumption $\mathrm{div}_x(b)^- \in L^1(0,T;L^\infty(\mathbb{T}^d))$.

At this point, we can conclude the proof by arguing as in Proposition 4.2.2 with Δt in place of δ. □

We are ready to prove Theorem 5.2.1, whose proof is now trivial:

Proof of Theorem 5.2.1. Combine Proposition 5.2.1 and Proposition 5.2.3 if $1 < p \leq d$. If $p > d$, we already proved the result in Proposition 5.2.3. □

5.3 Singular probabilistic approximation

The goal of this section is to use the results from Section 5.2 to approximate, in a suitable sense that we shall specify shortly, the solution u of (PDE). The idea is to exploit the Lagrangian representation $u(t,\cdot) = \Phi(t,\cdot)_\# u_0$ of the solution of (PDE) by constructing a vector field Φ_E and a measure μ which approximate Φ and u_0 respectively, so that $u_E(t) := \Phi_E(t,\cdot)_\#\mu$ approximates u in some sense. Before stating the main result of this section, let us give the definition of Wasserstein distance. For simplicity, we state this definition and some related results in the metric space $(\mathbb{R}^d, \mathrm{d}_{\mathbb{R}^d})$ rather than $(\mathbb{T}^d, \mathrm{d}_{\mathbb{T}^d})$. The same results, however, also hold in the latter case.

Definition 5.3.1. ***Wasserstein distance***
Consider μ,ν two positive measures in $(\mathbb{R}^d, \mathrm{d}_{\mathbb{R}^d})$, such that $\mu(\mathbb{R}^d) = \nu(\mathbb{R}^d)$. We define the Wasserstein distance *between μ and ν as*

$$\mathrm{W}(\mu,\nu) := \inf_{\pi\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{d}_{\mathbb{R}^d}(x,y)d\pi(x,y) \tag{5.3.1}$$

where Γ is the set of measures on $\mathbb{R}^d\times\mathbb{R}^d$ whose marginals are μ and ν. If in particular $\mathrm{d}_{\mathbb{R}^d}(x,y) = |x-y|$, we denote with W_1 the 1-Wasserstein distance:

$$\mathrm{W}_1(\mu,\nu) := \inf_{\pi\in\Gamma(\mu,\nu)} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|d\pi(x,y). \tag{5.3.2}$$

We also define a logarithmic Wasserstein distance *with respect to the distance*

$$\mathrm{d}_\alpha(x,y) := \log\left(1 + \frac{|x-y|}{h^\alpha}\right) \quad \textit{for some } \alpha \in [0,1] \textit{ and } h > 0, \tag{5.3.3}$$

which we will denote by $\widetilde{\mathrm{W}}_\alpha$. We will never explicitly write the dependence of d_α and $\widetilde{\mathrm{W}}_\alpha$ on h in order to have a lighter notation.

We will not use the above definition, but rather its dual formulation given by the Kantorovich-Rubinstein duality theorem:

Theorem 5.3.1. ***Kantorovich-Rubinstein duality theorem, [157]***
Let μ, ν be positive measures with compact support on the metric space $(\mathbb{R}^d, \mathrm{d}_{\mathbb{R}^d})$ and with the same mass, i.e. such that $\mu(\mathbb{R}^d) = \nu(\mathbb{R}^d)$. Then, it holds that

$$\inf_{\pi \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathrm{d}_{\mathbb{R}^d}(x,y) d\pi(x,y) = \sup_{\mathrm{Lip}(f) \leq 1} \int_{\mathbb{R}^d} f(x) d(\mu - \nu)(x),$$

where f is 1-Lipschitz with respect to the distance $\mathrm{d}_{\mathbb{R}^d}$.

We can now state the two main results of this section, which we present in two separate theorems: one for the 1-Wasserstein distance and the other for the logarithmic one. The reason is that a single statement would have been too long, and while the proofs of both theorems are very similar we believe both of them deserve to be presented in detail. Indeed while on the one hand the 1-Wasserstein distance is far more commonly used than the logarithmic one, on the other hand the latter is more suited to precisely capture the rate of convergence of the method. We recall that the first time such a logarithmic distance was used (at least in the context of the continuity equation in the DiPerna-Lions setting) is in [144] and that in [142, 143] the authors use such a logarithmic Wasserstein distance to evaluate the rate of convergence of their approximation method, so we decided to use the same distance to have a clearer comparison. We also remark that, for the sake of simplicity, in this section we will restrict ourselves to the case $u_0 \in L^\infty(\mathbb{T}^d)$ and non-negative. The non-negativity assumption is clearly not really an issue: we can always split the initial datum in its positive and negative parts. We will explain how to deal with the unbounded case in Remark 5.3.5. As a final note, we point out that we are forced to distinguish between bounded and unbounded initial densities due to our choice of a uniform probability $\bar{\mathrm{P}}$, which will play a crucial role in Theorem 5.3.2 and Theorem 5.3.3. With a different choice of $\bar{\mathrm{P}}$ this issue can be fixed and one does not need to distinguish between the two cases. We refer to Remark 5.3.6 for a more detailed discussion on the matter. Let us now state the two main results of this section:

Theorem 5.3.2. ***Singular probabilistic approximation of u in 1-Wasserstein distance***
Assume that b satisfies (MA) *and $u_0 \geq 0$ and bounded, and let u be the solution of* (PDE). *Assume also that the domain $\mathbb{T}^d$ is partitioned into sets $\{Q_i\}_{i=1}^N$ such that $\mathrm{diam}(Q_i) \leq \Delta x$ $\forall i = 1, \dots, N$ for some space discretization scale $\Delta x > 0$. Choose, for every set Q_i, a random starting point $x_0^i \in Q_i$ according to a uniform probability distribution on Q_i, and define $M_i := \int_{Q_i} u_0 dx$. For $t \in [0,T]$ consider the measure given by:*

$$u_E(t, \bar{x}_0) := \sum_{i=1}^N M_i \delta_{\Phi_E(t, x_0^i)} = \Phi_E(t, \cdot)_\# \left(\sum_{i=1}^N M_i \delta_{x_0^i} \right), \tag{5.3.4}$$

where $\bar{x}_0 = (x_0^1, x_0^2, \dots, x_0^N)$ and Φ_E is defined as in (5.2.2) *with a time step $\Delta t > 0$ and with $\delta = \sqrt{\Delta t}$ if $1 < p \leq d$ or as in* (5.2.3) *otherwise. Denoting with W_1 the 1-Wasserstein distance defined in* (5.3.2) *it holds that, in expected value, for every $\alpha \in (0,1)$:*

$$\mathbb{E}[\mathrm{W}_1(u(t,\cdot), u_E(t, \bar{x}_0))] \lesssim C_t |\log(\Delta t)|^{-1} + (\Delta x)^{1-\alpha} + t\alpha^{-p} |\log(\Delta x)|^{-p}, \tag{5.3.5}$$

with $C_t \to 0$ as $t \to 0$. We can also estimate the variance of $\mathrm{W}_1(u(t,\cdot), u_E(t, \bar{x}_0))$ if $N \lesssim (\Delta x)^{-d}$. In this case:

$$\begin{aligned} \mathrm{Var}[\mathrm{W}_1(u(t,\cdot), u_E(t, \bar{x}_0))] \lesssim C_t^2 |\log(\Delta t)|^{-2} + (\Delta x)^{2-2\alpha} + tC_t |\log(\Delta t)|^{-1} \\ + t^2 \alpha^{-p} |\log(\Delta x)|^{-p}. \end{aligned} \tag{5.3.6}$$

Theorem 5.3.3. ***Singular probabilistic approximation of u in logarithmic Wasserstein distance***
Consider the same setting of Theorem 5.3.2. It holds that, given the Wasserstein distance $\widetilde{\mathrm{W}}_\alpha$ defined in Definition 5.3.1 with $h := \max\{\Delta t, \Delta x\}$ and assuming $\alpha = 1/2$, if $1 < p \leq d$:

$$\mathbb{E}[\widetilde{\mathrm{W}}_{1/2}(u(t,\cdot),u_E(t,\bar{x}_0))] \lesssim 1 \tag{5.3.7}$$

and if $N \lesssim (\Delta x)^{-d}$, we also can estimate the variance as

$$\mathrm{Var}[\widetilde{\mathrm{W}}_{1/2}(u(t,\cdot),u_E(t,\bar{x}_0))] \lesssim |\log(h)|. \tag{5.3.8}$$

If $p > d$, it holds that for any $\alpha \in (0,1)$:

$$\mathbb{E}[\widetilde{\mathrm{W}}_{1-\alpha}(u(t,\cdot),u_E(t,\bar{x}_0))] \lesssim 1 + \frac{|\log(h)|^{1-p}}{\alpha^p} \tag{5.3.9}$$

and if $N \lesssim (\Delta x)^{-d}$:

$$\mathrm{Var}[\widetilde{\mathrm{W}}_{1-\alpha}(u(t,\cdot),u_E(t,\bar{x}_0))] \lesssim |\log(h)| + \frac{|\log(h)|^{2-p}}{\alpha^p}. \tag{5.3.10}$$

If moreover $p \geq 2$, we have that:

$$\mathrm{Var}[\widetilde{\mathrm{W}}_{1/2}(u(t,\cdot),u_E(t,\bar{x}_0))] \lesssim 1 \tag{5.3.11}$$

if $1 < p \leq d$, and

$$\mathrm{Var}[\widetilde{\mathrm{W}}_{1-\alpha}(u(t,\cdot),u_E(t,\bar{x}_0))] \lesssim 1 + \frac{|\log(h)|^{2-p}}{\alpha^p} \tag{5.3.12}$$

if $p > d$.

Before proving Theorem 5.3.2 and Theorem 5.3.3, we will need the following simple lemma:

Lemma 5.3.1.
Let μ, ν be two positive measures with compact support on $\mathbb{R}^d$ and such that $\mu(\mathbb{R}^d) = \nu(\mathbb{R}^d)$. Let μ_i, ν_i with $i = 1, \dots, N$ be positive measures such that

$$\mu = \sum_{i=1}^N \mu_i, \quad \nu = \sum_{i=1}^N \nu_i \quad \text{and } M_i = \mu_i(\mathbb{R}^d) = \nu_i(\mathbb{R}^d) \text{ for every } i.$$

Define also the probability measures

$$\bar{\mu}_i := \mu_i / M_i \text{ and } \bar{\nu}_i := \nu_i / M_i.$$

Then

$$\mathrm{W}(\mu,\nu) \leq \sum_{i=1}^N \mathrm{W}(\mu_i,\nu_i) = \sum_{i=1}^N M_i \, \mathrm{W}(\bar{\mu}_i,\bar{\nu}_i). \tag{5.3.13}$$

Proof. First of all let us prove that for every i it holds

$$\mathrm{W}(\mu_i,\nu_i) = M \, \mathrm{W}_1(\bar{\mu}_i,\bar{\nu}_i).$$

By definition:

$$\mathrm{W}(\bar{\mu}_i,\bar{\nu}_i) = \inf_{\pi \in \Gamma(\bar{\mu}_i,\bar{\nu}_i)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathrm{d}_{\mathbb{R}^d}(x,y) d\pi(x,y)$$

and given a sequence of measures π_n realising the infimum in (5.3.1) with marginals $\bar{\mu}_i$ and $\bar{\nu}_i$, the measures $M_i \pi_n$ will realise the infimum in (5.3.1) with marginals μ_i, ν_i, so

$$\mathrm{W}_1(\mu_i,\nu_i) = M_i \, \mathrm{W}_1(\bar{\mu}_i,\bar{\nu}_i).$$

For every i and every $\varepsilon > 0$ we consider π_i a transport plan between $\bar{\mu}_i$ and $\bar{\nu}_i$ such that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \mathrm{d}_{\mathbb{R}^d}(x,y) d\pi_i(x,y) \leq \mathrm{W}(\bar{\mu}_i,\bar{\nu}_i) + \varepsilon / N.$$

It is easy to see that $\pi^\varepsilon := \sum_{i=1}^N M_i \pi_i \in \Gamma(\mu,\nu)$, since for every A, B Borel:

$$\pi^\varepsilon(A\times\mathbb{R}^d) = \sum_{i=1}^N M_i\pi_i(A\times\mathbb{R}^d) = \sum_{i=1}^N M_i\bar{\mu}_i(A) = \sum_{i=1}^N \mu_i(A) = \mu(A),$$

and similarly we see that $\pi^\varepsilon(\mathbb{R}^d\times B) = \nu(B)$. Then:

$$\begin{aligned} \mathrm{W}(\mu,\nu) &\le \int_{\mathbb{R}^d\times\mathbb{R}^d} \mathrm{d}_{\mathbb{R}^d}(x,y)d\pi^\varepsilon(x,y) = \sum_{i=1}^N M_i \int_{\mathbb{R}^d\times\mathbb{R}^d} \mathrm{d}_{\mathbb{R}^d}(x,y)d\pi_i(x,y) \\ &\le \sum_{i=1}^N M_i\,\mathrm{W}(\bar{\mu}_i,\bar{\nu}_i) + \varepsilon = \sum_{i=1}^N \mathrm{W}(\mu_i,\nu_i) + \varepsilon. \end{aligned}$$

We conclude since ε is arbitrarily small. □

Remark 5.3.1.
Lemma 5.3.1 has an heuristic obvious "proof". What we are doing is simply to impose that the mass of μ_i is sent to ν_i for every i, and we minimize the cost for every $i = 1,\dots,N$. Of course by doing this we are imposing more constraints on $\pi \in \Gamma(\mu,\nu)$, so the inequality is just a consequence of the fact that we are considering the infimum in (5.3.1) *on a smaller set than $\Gamma(\mu,\nu)$.*

We are now ready to prove the theorems.

Proof of Theorem 5.3.2. First of all, we notice that in order to prove (5.3.5) we can work without loss of generality with Φ_δ. By Proposition 5.2.1, using Kantorovich-Rubinstein duality, we have:

$$\begin{aligned} \mathrm{W}_1(\Phi(t,\cdot)_\# u_0, \Phi_\delta(t,\cdot)_\# u_0) &\le \int_{\mathbb{T}^d} u_0(x)|\Phi(t,x)-\Phi_\delta(t,x)|dx \\ &\le \|u_0\|_{L^q}\|\Phi(t,\cdot)-\Phi_\delta(t,\cdot)\|_{L^p} \lesssim C_t|\log(\delta)|^{-1} \approx C_t|\log(\Delta t)|^{-1}. \end{aligned} \tag{5.3.14}$$

At time t fixed, consider $\bar{x}_0 = (x_0^1,\dots,x_0^N)$ a random variable in $((\mathbb{T}^d)^N, (\mathcal{B}(\mathbb{T}^d))^N, \bar{P})$, with $\mathcal{B}(\mathbb{T}^d)$ the Borel sigma algebra on $\mathbb{T}^d$ and

$$\bar{\mathrm{P}} = \mathrm{P}_1\otimes\mathrm{P}_2\cdots\otimes\mathrm{P}_N, \text{ with } \mathrm{P}_i \text{ a uniform distribution over } Q_i \text{ for every } i. \tag{5.3.15}$$

We also define $u_0^i := u_0 \llcorner Q_i$ and $\bar{u}_0^i := (u_0\llcorner Q_i)/M_i$ if $M_i > 0$, otherwise $M_i = \bar{u}_0^i = 0$, where $\mu\llcorner A$ denotes the restriction of a measure μ to the set A. In the following, each time we consider an expected value $\mathbb{E}[\cdot]$ or a variance $\mathrm{Var}[\cdot]$ it will be with respect to the probability $\bar{\mathrm{P}}$. We consider

$$\mu_i := \Phi_\delta(t,\cdot)_\#(u_0\llcorner Q_i) \text{ and } \nu_i := M_i\delta_{\Phi_E(t,x_0^i)} = \Phi_E(t,\cdot)_\#(M_i\delta_{x_0^i}),$$

which have the same mass, so we can apply Lemma 5.3.1. For the moment we will work in a fixed set $Q = Q_i$ of the partition such that $M_i \neq 0$, so in order to keep the notation lighter we will drop the subscript and in the following we will implicitly assume that the measures are restricted to Q and normalized, we will also write x_0 instead of x_0^i. Since $u(t,\cdot) = \Phi_\delta(t,\cdot)_\#\bar{u}_0$, we need to estimate $\mathbb{E}[\mathrm{W}_1(\Phi_\delta(t,\cdot)_\#\bar{u}_0, \delta_{\Phi_E(t,x_0)})]$. By triangle inequality

$$\begin{aligned} \mathbb{E}[\mathrm{W}_1(\Phi_\delta(t,\cdot)_\#\bar{u}_0, \delta_{\Phi_E(t,x_0)})] \le& \mathbb{E}[\mathrm{W}_1(\Phi_\delta(t,\cdot)_\#\bar{u}_0, \delta_{\Phi_\delta(t,x_0)})] \\ &+ \mathbb{E}[\mathrm{W}_1(\delta_{\Phi_\delta(t,x_0)}, \delta_{\Phi_E(t,x_0)})]. \end{aligned} \tag{5.3.16}$$

For the first addendum in (5.3.16), recalling that $\int_Q \bar{u}_0(x)dx = 1$:

$$\begin{aligned} &W_1(\Phi_\delta(t,\cdot)_\# \bar{u}_0, \delta_{\Phi_\delta(t,x_0)}) \\ &= \sup_{\mathrm{Lip}(h)\leq 1} \left(\int_Q \bar{u}_0(x) h(\Phi_\delta(t,x))dx - \int_Q \bar{u}_0(x) h(\Phi_\delta(t,x_0))dx \right) \\ &\leq \int_Q \bar{u}_0(x) |\Phi_\delta(t,x) - \Phi_\delta(t,x_0)| dx. \end{aligned} \tag{5.3.17}$$

Using Proposition 4.2.1 we can consider a set $K \subset \mathbb{T}^d$ with $|K^c| = \alpha^{-p} 2^p A_p(R,\Phi_\delta)^p |\log(\Delta x)|^{-p}$ for $\alpha \in (0,1)$ such that $\mathrm{Lip}(\Phi_\delta(t,\cdot) \llcorner K) \leq (\Delta x)^{-\alpha}$ for every $0 \leq t \leq T$. Notice that $A_p(R,\Phi_\delta)$ does not depend on the mollification scale δ, since $\|D_x b_\delta\|_{L^1(L^p)} \leq \|D_x b\|_{L^1(L^p)}$ for every $\delta > 0$ by Young's convolution inequality. Then using that $|x - x_0| \leq \mathrm{diam}(Q) \leq \Delta x$ and (5.3.17):

$$\begin{aligned} &\mathbb{E}\left[W_1(\Phi_\delta(t,\cdot)_\# \bar{u}_0, \delta_{\Phi_\delta(t,x_0)}) \right] \\ &\leq \frac{1}{\mathscr{L}^d(Q)} \left[\int_{Q\cap K} \left(\int_{Q\cap K} \bar{u}_0(x) |\Phi_\delta(t,x) - \Phi_\delta(t,x_0)| dx \right) dx_0 \right. \\ &+ \int_{Q\cap K} \left(\int_{Q\cap K^c} \bar{u}_0(x) |\Phi_\delta(t,x) - \Phi_\delta(t,x_0)| dx \right) dx_0 \\ &+ \left. \int_{Q\cap K^c} \left(\int_{Q} \bar{u}_0(x) |\Phi_\delta(t,x) - \Phi_\delta(t,x_0)| dx \right) dx_0 \right] \\ &\lesssim (\Delta x)^{1-\alpha} \|\bar{u}_0\|_{L^1(Q\cap K)} + (\Delta x + 2t\|b\|_\infty) \left(\|\bar{u}_0\|_{L^1(Q\cap K^c)} + \frac{\mathscr{L}^d(Q\cap K^c)}{\mathscr{L}^d(Q)} \right), \end{aligned} \tag{5.3.18}$$

where we used that

$$|\Phi_\delta(t,x) - \Phi_E(t,x_0)| \leq \Delta x + 2t\|b\|_\infty \text{ and that } \|\bar{u}_0\|_{L^1(Q)} = 1 \text{ by construction.}$$

As for the other addendum in (5.3.16), we have:

$$\begin{aligned} W_1(\delta_{\Phi_\delta(t,x_0)}, \delta_{\Phi_E(t,x_0)}) &= \sup_{\mathrm{Lip}(h)=1} \left(\int_{\mathbb{T}^d} h(x) d\delta_{\Phi_\delta(t,x_0)} - \int_{\mathbb{T}^d} h(x) d\delta_{\Phi_E(t,x_0)} \right) \\ &\leq |\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|. \end{aligned} \tag{5.3.19}$$

So, considering the expected value:

$$\mathbb{E}[W_1(\delta_{\Phi_\delta(t,x_0)}, \delta_{\Phi_E(t,x_0)})] \leq \frac{1}{\mathscr{L}^d(Q)} \int_Q |\Phi_\delta(t,x_0) - \Phi_E(t,x_0)| dx_0. \tag{5.3.20}$$

By Lemma 5.3.1:

$$\begin{aligned} &\mathbb{E}[W_1(\Phi_\delta(t,\cdot)_\# u_0, u_E(t,\bar{x}_0))] \leq \mathbb{E}\left[\sum_{i=1}^N M_i W_1\left(\Phi_\delta(t,\cdot)_\# \bar{u}_0^i, \delta_{\Phi_E(t,x_0^i)} \right) \right] \\ &= \sum_{i=1}^N M_i \mathbb{E}\left[W_1\left(\Phi_\delta(t,\cdot)_\# \bar{u}_0^i, \delta_{\Phi_E(t,x_0^i)} \right) \right] \\ &\lesssim \sum_{i=1}^N M_i \left((\Delta x)^{1-\alpha} \|\bar{u}_0^i\|_{L^1(Q_i\cap K)} + (\Delta x + 2t\|b\|_\infty) \|\bar{u}_0^i\|_{L^1(Q_i\cap K^c)} \right. \\ &\qquad \left. + (\Delta x + 2t\|b\|_\infty) \frac{\mathscr{L}^d(Q\cap K^c)}{\mathscr{L}^d(Q)} + \frac{1}{\mathscr{L}^d(Q_i)} \int_{Q_i} |\Phi_E(t,x_0) - \Phi_\delta(t,x_0)| dx_0 \right) \end{aligned} \tag{5.3.21}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} \Bigg((\Delta x)^{1-\alpha} \|u_0^i\|_{L^1(Q_i\cap K)} + (\Delta x + 2t\|b\|_\infty)\|u_0^i\|_{L^1(Q_i\cap K^c)} \\
&\qquad + \|u_0\|_\infty (\Delta x + 2t\|b\|_\infty)\mathscr{L}^d(Q_i\cap K^c) + \|u_0\|_\infty \int_{Q_i} |\Phi_E(t,x_0) - \Phi_\delta(t,x_0)|dx_0 \Bigg) \\
&\lesssim (\Delta x)^{1-\alpha}\|u_0\|_{L^1(\mathbb{T}^d)} + t\|u_0\|_{L^1(K^c)} + t\|u_0\|_\infty \mathscr{L}^d(K^c) + \|u_0\|_\infty \|\Phi_\delta(t,\cdot) - \Phi_E(t,\cdot)\|_{L^1(\mathbb{T}^d)} \\
&\lesssim (\Delta x)^{1-\alpha} + \|u_0\|_\infty \left(t\frac{|\log(\Delta x)|^{-p}}{\alpha^p} + C_t |\log(\Delta t)|^{-1} \right),
\end{aligned}
$$

where we use that $\|u_0\|_{L^1(K^c)} \le \|u_0\|_\infty \mathscr{L}^d(K^c)$ with $\mathscr{L}^d(K^c) \approx \alpha^{-p}|\log(\Delta x)|^{-p}$. We also have

$$
\|\Phi_\delta(t,\cdot) - \Phi_E(t,\cdot)\|_{L^1(\mathbb{T}^d)} \lesssim C_t |\log(\Delta t)|^{-1} \tag{5.3.22}
$$

by Hölder's inequality and Proposition 5.2.3, and thus we have proved (5.3.5). We now estimate the variance. By definition

$$
\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.
$$

We will simply estimate $\mathrm{Var}(X)$ from above with $\mathbb{E}[X^2]$. First of all we notice that, in order to prove (5.3.6), we can work without loss of generality with Φ_δ instead of Φ. Indeed:

$$
\mathrm{W}_1(\Phi(t,\cdot)_\# \bar u_0, u_E(t,\bar x_0))^2 \le 2\,\mathrm{W}_1(\Phi(t,\cdot)_\# \bar u_0, \Phi_\delta(t,\cdot)_\# \bar u_0)^2 + 2\,\mathrm{W}_1(\Phi_\delta(t,\cdot)_\# \bar u_0, u_E(t,\bar x_0))^2, \tag{5.3.23}
$$

and by (5.3.14) we have

$$
\mathrm{W}_1(\Phi(t,\cdot)_\# \bar u_0, \Phi_\delta(t,\cdot)_\# \bar u_0)^2 \lesssim C_t^2 |\log(\Delta t)|^{-2}. \tag{5.3.24}
$$

We now restrict to working in a fixed set Q_i. As we did before, we omit writing the index i and we will denote $(u_0 \llcorner Q_i)/M_i$ and x_0^i as $\bar u_0$ and x_0. It holds:

$$
\mathrm{W}_1(\Phi_\delta(t,\cdot)_\# \bar u_0, \delta_{\Phi_E(t,x_0)})^2 \le 2\,\mathrm{W}_1(\Phi_\delta(t,\cdot)_\# \bar u_0, \delta_{\Phi_\delta(t,x_0)})^2 + 2\,\mathrm{W}_1(\delta_{\Phi_\delta(t,x_0)}, \delta_{\Phi_E(t,x_0)})^2. \tag{5.3.25}
$$

For the first addendum in the right-hand side of (5.3.25), using (5.3.17), the same choice of K, the decomposition in (5.3.18) and the elementary inequality $(a+b+c)^2 \le 4a^2 + 4b^2 + 4c^2$:

$$
\begin{aligned}
\mathbb{E}[\mathrm{W}_1(\Phi_\delta(t,\cdot)_\# \bar u_0, \delta_{\Phi_\delta(t,x_0)})^2] \lesssim (\Delta x)^{2-2\alpha}\|\bar u_0\|^2_{L^1(Q\cap K)} &+ (\Delta x + 2\|b\|_\infty t)^2 \|\bar u_0\|^2_{L^1(Q\cap K^c)} \\
&+ (\Delta x + 2\|b\|_\infty t)^2 \frac{\mathscr{L}^d(Q\cap K^c)^2}{\mathscr{L}(Q)^2}.
\end{aligned} \tag{5.3.26}
$$

For the second addendum in (5.3.25), using (5.3.19) in expected value:

$$
\begin{aligned}
\mathbb{E}[\mathrm{W}_1(\delta_{\Phi_\delta(t,x_0)}, \delta_{\Phi_E(t,x_0)})^2] &\le \frac{1}{\mathscr{L}^d(Q)} \int_Q |\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|^2 dx_0 \\
&\le \frac{\|\Phi_\delta(t,\cdot) - \Phi_E(t,\cdot)\|_\infty}{\mathscr{L}^d(Q)} \int_Q |\Phi_\delta(t,x_0) - \Phi_E(t,x_0)| dx_0 \\
&\le \frac{2\|b\|_\infty t}{\mathscr{L}^d(Q)} \int_Q |\Phi_\delta(t,x_0) - \Phi_E(t,x_0)| dx_0.
\end{aligned} \tag{5.3.27}
$$

The conclusion follows as in (5.3.21), using (5.3.26), (5.3.27) and that by Cauchy-Schwartz:

$$
\mathrm{W}_1(\mu,\nu) \le \sum_{i=1}^{N} M_i\,\mathrm{W}_1(\bar\mu_i, \bar\nu_i) \implies \mathrm{W}_1(\mu,\nu)^2 \le N \sum_{i=1}^{N} M_i^2\,\mathrm{W}_1(\bar\mu_i,\bar\nu_i)^2.
$$

Indeed

$$
\begin{aligned}
\mathbb{E}[\mathrm{W}_1(\Phi(t,\cdot)_\# u_0, u_E(t,\bar{x}_0))^2] &\leq N\sum_{i=1}^N M_i^2 \mathbb{E}\left[\mathrm{W}_1\left(\Phi(t,\cdot)_\#\bar{u}_0^i, \delta_{\Phi_E(t,x_0^i)}\right)^2\right] \qquad (5.3.28)\\
&\lesssim N\sum_{i=1}^N M_i^2 \Bigg((\Delta x)^{2-2\alpha}\|\bar{u}_0^i\|^2_{L^1(Q_i\cap K)} + (\Delta x+t)^2\|\bar{u}_0^i\|^2_{L^1(Q_i\cap K^c)}\\
&\qquad + (\Delta x+t)^2\frac{\mathscr{L}^d(Q_i\cap K^c)^2}{\mathscr{L}^d(Q_i)^2} + \frac{t}{\mathscr{L}^d(Q_i)}\int_{Q_i}|\Phi_E(t,x_0)-\Phi_\delta(t,x_0)|dx_0\Bigg)\\
&\leq N\sum_{i=1}^N M_i \Bigg((\Delta x)^{2-2\alpha}\|u_0^i\|_{L^1(Q_i)} + (2(\Delta x)^2+2t^2)\|u_0^i\|_{L^1(Q_i\cap K^c)}\\
&\qquad + (2(\Delta x)^2+2t^2)\|u_0\|_\infty \mathscr{L}(K^c\cap Q_i) + t\|u_0\|_\infty\int_{Q_i}|\Phi_E(t,x_0)-\Phi_\delta(t,x_0)|dx_0\Bigg)\\
&\leq N\sum_{i=1}^N M_i \Bigg((\Delta x)^{2-2\alpha}\|u_0^i\|_{L^1(Q_i)} + (4(\Delta x)^2+4t^2)\|u_0\|_\infty \mathscr{L}(K^c\cap Q_i)\\
&\qquad + t\|u_0\|_\infty\int_{Q_i}|\Phi_E(t,x_0)-\Phi_\delta(t,x_0)|dx_0\Bigg).
\end{aligned}
$$

By the assumption on $\mathrm{diam}(Q_i)$ we have $M_i \lesssim \|u_0\|_\infty(\Delta x)^d$ for every i, and also by assumption $N \lesssim (\Delta x)^{-d}$, so

$$
\begin{aligned}
&\mathbb{E}[\mathrm{W}_1(\Phi_\delta(t,\cdot)_\# u_0, u_E(t,\bar{x}_0))^2]\\
&\lesssim \|u_0\|_\infty \sum_{i=1}^N \Bigg((\Delta x)^{2-2\alpha}\|u_0^i\|_{L^1(Q_i)} + t^2\|u_0\|_\infty \mathscr{L}^d(K^c\cap Q_i)\\
&\qquad + t\|u_0\|_{L^\infty}\int_{Q_i}|\Phi_E(t,x_0)-\Phi_\delta(t,x_0)|dx_0\Bigg)\\
&= \|u_0\|_\infty \Bigg((\Delta x)^{2-2\alpha}\|u_0\|_{L^1(\mathbb{T}^d)} + t^2\|u_0\|_\infty \mathscr{L}^d(K^c)\\
&\qquad + t\|u_0\|_\infty\int_{\mathbb{T}^d}|\Phi_E(t,x_0)-\Phi_\delta(t,x_0)|dx_0\Bigg)\\
&\leq \|u_0\|_\infty \Bigg((\Delta x)^{2-2\alpha}\|u_0\|_{L^1(\mathbb{T}^d)} + t^2\|u_0\|_\infty \mathscr{L}^d(K^c)\\
&\qquad + t\|u_0\|_\infty\int_{\mathbb{T}^d}|\Phi_E(t,x_0)-\Phi_\delta(t,x_0)|dx_0\Bigg)\\
&\lesssim \|u_0\|_\infty \left((\Delta x)^{2-2\alpha}\|u_0\|_{L^1(\mathbb{T}^d)} + t^2\|u_0\|_\infty \alpha^{-p}|\log(\Delta x)|^{-p} + tC_t\|u_0\|_\infty|\log(\Delta t)|^{-1}\right),
\end{aligned}
$$

and together with (5.3.23) and (5.3.24) we can conclude and get (5.3.6). □

Remark 5.3.2.
A simple sufficient condition to impose in order to have that $N \lesssim (\Delta x)^{-d}$ is a lower bound on the volume of the cells, i.e. $|Q_i| \geq c(\Delta x)^d$ for some $c > 0$ for every i.

Remark 5.3.3. ***The role of*** α
We will try to give an heuristic explanation of the role of α. It essentially takes into account the fact that one cannot estimate $u(t,\cdot)$ better than a certain threshold even for small times, since we have a discretization of the initial datum at scale Δx. We can check that, if $t = 0$, we can consider $\alpha = 1$ and estimate $\mathrm{W}_1\left(u_0, \sum_{i=1}^N M_i\delta_{x_0^i}\right)$ deterministically for every choice of $x_0^i \in Q_i$. Indeed, consider $u_0^i = u_0 \llcorner Q_i$. By Kantorovich-Rubinstein duality:

$$\mathrm{W}_1(u_0^i, M_i\delta_{x_0^i}) = \sup_{\mathrm{Lip}(f)\le 1}\left(\int_{Q_i} f(x)u_0^i(x)dx - \int_{Q_i} f(x_0)u_0^i(x)dx\right)$$
$$= \int_{Q_i} u_0^i(x)|f(x)-f(x_0)|dx \le \Delta x \int_{Q_i} u_0^i(x)dx,$$

and we can easily conclude applying Lemma 5.3.1.

Remark 5.3.4. ***The role of the expected value***
If we did not consider any expected value in Theorem 5.3.2, instead of (5.3.20) *we would have had to estimate pointwise differences between Φ_δ and Φ_E, which we are unable to control.*

Before proving Theorem 5.3.3, let us remark once again why it is more precise than Theorem 5.3.2. While in Theorem 5.3.2 we have, for every $\alpha \in (0,1)$, essentially the same logarithmic rate, in Theorem 5.3.3 by defining $h := \max\{\Delta t, \Delta x\}$ the optimal rate h^α for which we will be able to estimate both the expected value and the variance will be $1/2$ for $1 < p \le d$, while for $p > d$ we get better estimates. Namely, if $p > d$ we can estimate the $\widetilde{\mathrm{W}}_{1-\alpha}$ distance, but with constants that diverge to $+\infty$ as $\alpha \to 0$. We thus do not have a sharp rate for this case. We notice that the convergence rate of $1/2$ is likely sharp if $1 < p \le d$ (as shown in [142] for the upwind scheme).

Proof of Theorem 5.3.3. Since the relevant quantity will be $h = \max\{\Delta t, \Delta x\}$ in the following, we can assume without loss of generality that $h = \Delta t = \Delta x$. If $1 < p \le d$, we first need to estimate the distance $\widetilde{\mathrm{W}}_{1/2}(\Phi(t,\cdot)_\# u_0, \Phi_\delta(t,\cdot)_\# u_0)$ since we want to work with Φ_δ in place of Φ with $\delta = \sqrt{h}$. Notice that this will not be necessary if $p > d$. By Kantorovich-Rubinstein duality

$$\widetilde{\mathrm{W}}_{1/2}(\Phi(t,\cdot)_\# u_0, \Phi_\delta(t,\cdot)_\# u_0) \le \int_{\mathbb{T}^d} u_0(x)\log\left(1 + \frac{|\Phi(t,x)-\Phi_\delta(t,x)|}{\sqrt{h}}\right)dx \tag{5.3.29}$$
$$\le \|u_0\|_{L^q(\mathbb{T}^d)}\left\|\log\left(1+\frac{|\Phi(t,\cdot)-\Phi_\delta(t,\cdot)|}{\sqrt{h}}\right)\right\|_{L^p(\mathbb{T}^d)} \le C,$$

where we have used the estimate of Remark 4.2.2, since $\|b - b_\delta\|_{L^1(L^p)} \lesssim \sqrt{h}$, because $\delta = \sqrt{h}$. Arguing as in Theorem 5.3.2 we restrict to working in a fixed set Q_i and we will omit the index for the moment to lighten the notation, in particular $\bar{u}_0$, x_0 and Q will denote $(u_0 \llcorner Q_i)/M_i$, x_0^i and Q_i respectively. By triangle inequality:

$$\mathbb{E}\left[\widetilde{\mathrm{W}}_{1/2}(\Phi_\delta(t,\cdot)_\#\bar{u}_0, \delta_{\Phi_E(t,x_0)})\right] \le \mathbb{E}\left[\widetilde{\mathrm{W}}_{1/2}(\Phi_\delta(t,\cdot)_\#\bar{u}_0, \delta_{\Phi_\delta(t,x_0)})\right] \tag{5.3.30}$$
$$+\, \mathbb{E}\left[\widetilde{\mathrm{W}}_{1/2}(\delta_{\Phi_\delta(t,x_0)}, \delta_{\Phi_E(t,x_0)})\right].$$

For the first addendum in the right-hand side of (5.3.30) it holds the analogous of (5.3.17), i.e.

$$\widetilde{\mathrm{W}}_{1/2}(\Phi_\delta(t,\cdot)_\#\bar{u}_0, \delta_{\Phi_\delta(t,x_0)}) \le \int_Q \bar{u}_0(x)\log\left(1+\frac{|\Phi_\delta(t,x)-\Phi_\delta(t,x_0)|}{\sqrt{h}}\right)dx.$$

By Proposition 4.2.1 we can choose $K \subset \mathbb{T}^d$ such that $|\mathbb{T}^d \setminus K| = 4^p A_p(R,\Phi_\delta)^p |\log(h)|^{-p}$ and so that Φ_δ satisfies $\mathrm{Lip}(\Phi_\delta(t,\cdot)\llcorner K) \le 1/\sqrt{h}$ for every t. Notice that $A_p(R,\Phi_\delta)$ does not depend on the mollification scale δ, since $\|D_x b_\delta\|_{L^1(L^p)} \le \|D_x b\|_{L^1(L^p)}$ for every $\delta > 0$ by Young's convolution inequality. Then it holds

$$\mathbb{E}[\widetilde{\mathrm{W}}_{1/2}(\Phi_\delta(t,\cdot)_\#\bar{u}_0, \delta_{\Phi_\delta(t,x_0)})]$$

$$
\begin{aligned}
&\leq \frac{1}{\mathscr{L}^d(Q)}\Bigg[\int_{Q\cap K}\left(\int_{Q\cap K}\bar{u}_0(x)\log\left(1+\frac{|\Phi_\delta(t,x)-\Phi_\delta(t,x_0)|}{\sqrt{h}}\right)dx\right)dx_0\\
&+\int_{Q\cap K}\left(\int_{Q\cap K^c}\bar{u}_0(x)\log\left(1+\frac{|\Phi_\delta(t,x)-\Phi_\delta(t,x_0)|}{\sqrt{h}}\right)dx\right)dx_0\\
&+\int_{Q\cap K^c}\left(\int_{Q}\bar{u}_0(x)\log\left(1+\frac{|\Phi_\delta(t,x)-\Phi_\delta(t,x_0)|}{\sqrt{h}}\right)dx\right)dx_0\Bigg]\\
&\lesssim \|\bar{u}_0\|_{L^1(Q\cap K)}+|\log(h)|\,\|\bar{u}_0\|_{L^1(Q\cap K^c)}+|\log(h)|\frac{\mathscr{L}^d(Q\cap K^c)}{\mathscr{L}^d(Q)},
\end{aligned}
$$

where in the first addendum we used the Lipschitz estimate on $\Phi_\delta(t,\cdot)\llcorner K$, while for the second and third addenda we simply estimated the logarithm as $|\log(h)|$ and used that $\|\bar{u}_0\|_{L^1(Q)}=1$. For the second addendum in (5.3.30), arguing as for (5.3.19):

$$
\mathbb{E}[\widetilde{\mathrm{W}}_{1/2}(\delta_{\Phi(t,x_0)},\delta_{\Phi_E(t,x_0)})]\leq\frac{1}{\mathscr{L}^d(Q)}\int_Q\log\left(1+\frac{|\Phi_\delta(t,x_0)-\Phi_E(t,x_0)|}{\sqrt{h}}\right)dx_0.
$$

So, estimating as in (5.3.21):

$$
\begin{aligned}
&\mathbb{E}[\widetilde{\mathrm{W}}_{1/2}(\Phi_\delta(t,\cdot)_\# u_0,u_E(t,\bar{x}_0))]\leq\mathbb{E}\left[\sum_{i=1}^N M_i\widetilde{\mathrm{W}}_{1/2}\left(\Phi_\delta(t,\cdot)_\#\bar{u}_0^i,\delta_{\Phi_E(t,x_0^i)}\right)\right] \qquad (5.3.31)\\
&=\sum_{i=1}^N M_i\mathbb{E}\left[\widetilde{\mathrm{W}}_{1/2}\left(\Phi_\delta(t,\cdot)_\#\bar{u}_0^i,\delta_{\Phi_E(t,x_0^i)}\right)\right]\\
&\lesssim\sum_{i=1}^N M_i\Bigg(\|\bar{u}_0^i\|_{L^1(Q_i\cap K)}+|\log(h)|\,\|\bar{u}_0^i\|_{L^1(Q_i\cap K^c)}\\
&\qquad+|\log(h)|\frac{\mathscr{L}^d(Q_i\cap K)}{\mathscr{L}^d(Q)}+\frac{1}{\mathscr{L}^d(Q_i)}\int_{Q_i}\log\left(1+\frac{|\Phi_\delta(t,x_0)-\Phi_E(t,x_0)|}{\sqrt{h}}\right)dx_0\Bigg)\\
&\leq\sum_{i=1}^N\Bigg(\|u_0^i\|_{L^1(Q_i\cap K)}+|\log(h)|\,\|u_0\|_\infty\mathscr{L}^d(Q_i\cap K^c)\\
&\qquad+|\log(h)|\,\|u_0\|_\infty\mathscr{L}^d(Q_i\cap K^c)+\|u_0\|_\infty\int_{Q_i}\log\left(1+\frac{|\Phi_\delta(t,x_0)-\Phi_E(t,x_0)|}{\sqrt{h}}\right)dx_0\Bigg)\\
&\lesssim\|u_0\|_{L^1(\mathbb{T}^d)}+|\log(h)|\,\|u_0\|_\infty\,\mathscr{L}^d(K^c)+\|u_0\|_\infty\int_{\mathbb{T}^d}\log\left(1+\frac{|\Phi_\delta(t,x_0)-\Phi_E(t,x_0)|}{\sqrt{h}}\right)dx_0\\
&\lesssim\|u_0\|_{L^1(\mathbb{T}^d)}+\|u_0\|_\infty\left(|\log(h)|^{1-p}+1\right)\leq C.
\end{aligned}
$$

Combining (5.3.29) and (5.3.31) we get (5.3.7). With the same estimates used in Theorem 5.3.2, we can also recover the estimate for the variance. First of all, let us see that we can work without loss of generality with Φ_δ in place of Φ. Indeed we can use (5.3.23) and (5.3.29) with $\widetilde{\mathrm{W}}_{1/2}$ to get $\widetilde{\mathrm{W}}_{1/2}(\Phi(t,\cdot)_\#\bar{u}_0,\Phi_\delta(t,\cdot)_\#\bar{u}_0)^2\leq C$. We are left to estimate $\widetilde{\mathrm{W}}_{1/2}(\Phi_\delta(t,\cdot)_\# u_0,u_E(t,\bar{x}_0))$. By (5.3.25) we first estimate:

$$
\begin{aligned}
&\mathbb{E}[\widetilde{\mathrm{W}}_{1/2}(\Phi_\delta(t,\cdot)_\#\bar{u}_0,\delta_{\Phi_\delta(t,x_0)})^2]\\
&\lesssim\frac{1}{\mathscr{L}^d(Q)}\Bigg[\int_{Q\cap K}\left(\int_{Q\cap K}\bar{u}_0(x)\log\left(1+\frac{|\Phi_\delta(t,x)-\Phi_\delta(t,x_0)|}{\sqrt{h}}\right)dx\right)^2dx_0\\
&+\int_{Q\cap K}\left(\int_{Q\cap K^c}\bar{u}_0(x)\log\left(1+\frac{|\Phi_\delta(t,x)-\Phi_\delta(t,x_0)|}{\sqrt{h}}\right)dx\right)^2dx_0
\end{aligned}
$$

$$+ \int_{Q\cap K^c} \left(\int_Q \bar{u}_0(x) \log\left(1 + \frac{|\Phi_\delta(t,x) - \Phi_\delta(t,x_0)|}{\sqrt{h}} \right) dx \right)^2 dx_0 \Bigg]$$
$$\lesssim \|\bar{u}_0\|_{L^1(Q\cap K)} + |\log(h)|^2 \, \|\bar{u}_0\|_{L^1(Q\cap K^c)} + |\log(h)|^2 \frac{\mathscr{L}^d(Q\cap K^c)}{\mathscr{L}^d(Q)},$$

where in the first addendum we used the Lipschitz estimate on $\Phi_\delta(t,\cdot) \llcorner K$ and in the last two addenda we estimated the logarithm as $|\log(h)|$. The other relevant estimate is:

$$\mathbb{E}\left[\widetilde{\mathrm{W}}_{1/2}(\delta_{\Phi_\delta(t,x_0)}, \delta_{\Phi_E(t,x_0)})^2 \right] \le \frac{1}{\mathscr{L}^d(Q)} \int_Q \log\left(1 + \frac{|\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|}{\sqrt{h}} \right)^2 dx_0 \tag{5.3.32}$$
$$\lesssim \frac{|\log(h)|}{\mathscr{L}^d(Q)} \int_Q \log\left(1 + \frac{|\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|}{\sqrt{h}} \right) dx_0.$$

Summing over j as in (5.3.28) we have

$$\mathrm{Var}[\widetilde{\mathrm{W}}_{1/2}(u(t,\cdot), u_E(t,\bar{x}_0))] \lesssim \|u_0\|_\infty \Bigg[\|u_0\|_{L^1(\mathbb{T}^d)} + |\log(h)|^2 \|u_0\|_\infty \mathscr{L}^d(K^c)$$
$$+ |\log(h)| \, \|u_0\|_\infty \int_{\mathbb{T}^d} \log\left(1 + \frac{|\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|}{\sqrt{h}} \right) dx_0 \Bigg]$$
$$\lesssim \|u_0\|_\infty (\|u_0\|_{L^1(\mathbb{T}^d)} + \|u_0\|_\infty \, |\log(h)|^{2-p} + \|u_0\|_\infty \, |\log(h)|)$$
$$\lesssim |\log(h)|,$$

so we have proved (5.3.8). The stronger rate for $p > d$ can be proved by noticing that we do not need (5.3.29), since we work directly with Φ, and by defining K such that

$$|\mathbb{T}^d \setminus K| \le \alpha^{-p} 2^p A_p(R,\Phi)^p |\log(h)|^{-p} \text{ and } \mathrm{Lip}(\Phi(t,\cdot) \llcorner K) \le h^{-\alpha} \text{ for every } t.$$

Every estimate works in the same way, but keeping track of the dependence on α gives us (5.3.9) and (5.3.10). Finally, if we also assume $p \ge 2$ we can keep the square in the integrand on the first line of (5.3.32), and with the same estimates as above we can conclude that:

$$\mathrm{Var}[\widetilde{\mathrm{W}}_{1/2}(u(t,\cdot), u_E(t,\bar{x}_0))] \lesssim \|u_0\|_{L^1(\mathbb{T}^d)} + \|u_0\|_{L^\infty} |\log(h)|^{2-p}$$
$$+ \|u_0\|_\infty \int_{\mathbb{T}^d} \log\left(1 + \frac{|\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|}{\sqrt{h}} \right)^2 dx_0.$$

Since $p \ge 2$ we can estimate the last integral using Hölder's inequality and (5.2.5), proving (5.3.11). Estimate (5.3.12) can be proved similarly, keeping track of the dependence of α on $\mathscr{L}^d(K^c)$. □

Remark 5.3.5. *The case* $u_0 \notin L^\infty(\mathbb{T}^d)$
As already remarked, for the sake of simplicity we assumed u_0 to be bounded. In u_0 is unbounded, one can simply estimate u_0 with $u_0 \mathbb{1}_{\{u_0 \le K\}}$, with $\mathbb{1}_A$ denoting the indicator function of the set A. After this first approximation one should keep track of the dependence of the estimates on K.

Remark 5.3.6. *A possible alternative choice for* $\bar{P}$
If the choice of the point x_0^i in the cell Q_i is not made according to a uniform distribution, but according to the probability density function $(u_0 \llcorner Q_i)/M_i$, we can verify that both Theorem 5.3.2 and Theorem 5.3.3 hold without the need to distinguish between the bounded and the unbounded cases. In particular, we point out that the crucial technical advantage lies in the fact that if x_0 is picked according to the ad hoc probability just defined, it holds:

$$M_i \mathbb{E}[\widetilde{\mathrm{W}}_{1/2}(\delta_{\Phi(t,x_0)}, \delta_{\Phi_E(t,x_0)})] = M_i \int_{Q_i} \bar{u}_0(x_0) \log\left(1 + \frac{|\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|}{\sqrt{h}}\right) dx_0 \tag{5.3.33}$$

$$= \int_{Q_i} u_0(x_0) \log\left(1 + \frac{|\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|}{\sqrt{h}}\right) dx_0,$$

while using a uniform distribution we needed to use that $M_i/\mathscr{L}^d(Q_i) \le \|u_0\|_\infty$, crucially relying on the boundedness of u_0. There are two reasons why we preferred to give the proofs using a uniform probability distribution even though it produces less sharp results. The first one is that a uniform probability distribution is simpler, and the second reason is that, in many practical cases, the initial distribution is bounded and the issue is given by the low regularity of the velocity field. We also notice that while using a bounded density we essentially only considered L^1 stability estimates for the flow, if we had used the ad hoc probability distribution just defined we would have really needed the full L^p stability estimate on the flows. Indeed after summing the terms like (5.3.33) *for every i we would have needed an estimate like:*

$$\int_{\mathbb{T}^d} u_0(x_0) \log\left(1 + \frac{|\Phi_\delta(t,x_0) - \Phi_E(t,x_0)|}{\sqrt{h}}\right) dx_0 \le$$

$$\|u_0\|_{L^q(\mathbb{T}^d)} \left\| \log\left(1 + \frac{|\Phi_\delta(t,\cdot) - \Phi_E(t,\cdot)|}{\sqrt{h}}\right) \right\|_{L^p(\mathbb{T}^d)}$$

to conclude.

Thanks to the above results we can now prove the following proposition, which in a sort of Monte Carlo-like method for approximating in Wasserstein distance the solution u of (PDE). For the sake of simplicity and since the arguments would be identical, we give the result for a specific choice of Δt, Δx, α and only with respect to the 1-Wasserstein distance.

Proposition 5.3.1. ***Monte Carlo estimate of solution of*** (PDE)
Let b, u_0 and $\{Q_i\}_{i=1}^N$ satisfy the assumptions of Theorem 5.3.2. Moreover, assume $\alpha = 1/2$, $\Delta t = \Delta x = h$ for some $h > 0$ and that $C_t|\log(h)|^{-1} \gg \sqrt{h}$. Then, defining

$$S_n(t) := \frac{1}{n}\sum_{i=1}^n u_E^i(t)$$

the mean of n realisations of u_E (see Definition 5.3.4; we omit to explicitly write in every u_E^i the dependence on $\bar{x}_0$) it holds that for every $k > 1$

$$\bar{\mathrm{P}}\left(|\,\mathrm{W}_1(S_n(t), u(t,\cdot)) - \mathbb{E}[S_n(t)]| > k\sqrt{\frac{C_t|\log(h)|^{-1}}{n}} \right) \le \frac{1}{k^2},$$

where $C_t \to 0$ as $t \to 0$ and $\mathbb{E}[\mathrm{W}_1(S_n(t), u(t,\cdot))] \le C_t|\log(h)|^{-1}$.

Proof. Let $X_i := \mathrm{W}_1(u(t,\cdot), u_E^i(t))$ be a random variable with respect to the probability $\bar{P}$ defined in (5.3.15). By Lemma 5.3.1:

$$\mathrm{W}_1(S_n(t), u(t,\cdot)) \le \frac{1}{n}\sum_{i=1}^n \mathrm{W}_1(u_E^i(t), u(t,\cdot)),$$

so by (5.3.5)

$$\mathbb{E}[\mathrm{W}_1(S_n(t), u(t,\cdot))] \le \frac{1}{n}\sum_{i=1}^n \mathbb{E}[\mathrm{W}_1(u_E^i(t), u(t,\cdot))] \lesssim C_t|\log(h)|^{-1}, \tag{5.3.34}$$

and we have an upper bound for $\mathbb{E}[\mathrm{W}_1(S_n(t), u(t,\cdot))]$. By construction the random variables X_i's are independent and identically distributed, so by (5.3.6)

$$\mathrm{Var}[S_n(t)] = \frac{\mathrm{Var}[X_1]}{n} \leq \frac{C_t|\log(h)|^{-1}}{n}. \tag{5.3.35}$$

Finally by Chebyshev's inequality and (5.3.35):

$$\begin{aligned}\bar{\mathrm{P}}\left(\left\{|S_n(t) - \mathbb{E}[S_n(t)]| > k\sqrt{\frac{C_t|\log(h)|^{-1}}{n}}\right\}\right) &\leq \bar{\mathrm{P}}\left(\left\{|S_n(t) - \mathbb{E}[S_n(t)]| > k\sqrt{\mathrm{Var}[S_n(t)]}\right\}\right)\\ &\leq \frac{1}{k^2}.\end{aligned}$$

□

5.4 Diffuse deterministic approximation

In the previous section we showed how, in expected value, an approximation of the initial datum by a sum of Dirac deltas advected by the vector field Φ_E led to an approximation of the density u which solves (PDE). The point was to prove that the analogue of the explicit Euler method applied to one of the simplest (weak) approximation of the initial datum (i.e. a sum of Dirac deltas) works also in this low-regularity setting. Such an approximation, however, is not well-suited for many applications where one really wants a diffuse approximation of the density rather than a singular one (for instance when there is a coupling between the continuity equation and other equations). Moreover, we want to also present an approximation scheme which works deterministically. We thus consider a variation of the previous results to cover also this important case, showing how allowing for a stronger approximation of the initial datum and a suitable approximation of the regular Lagrangian flow Φ we can give a deterministic approximation. As we have just pointed out there are two major problems with the method presented in Section 5.3:

1. The approximation of u is a singular measure. Often, one would like a diffuse approximation of a diffuse initial datum.

2. The estimates are valid only in expected value, essentially due to the high sensitivity (at least a priori, see Remark 5.3.4) to the choice of the points $x_0^i \in Q_i$.

A way to fix the above issues is to approximate the flow Φ by performing an additional averaging in space and to consider a strong (in L^1 norm) approximation of the initial datum. In this way, deterministic estimates for both the 1-Wasserstein distance and for the logarithmic Wasserstein distance introduced in the previous section will follow easily, as we will see in Theorem 5.4.1. Let us now make these ideas rigorous.

Definition 5.4.1. ***Definition of*** $\bar{\Phi}_E$
Assume the domain $\mathbb{T}^d$ *is partitioned in sets* $\{Q_i\}_{i=1}^n$. *Let* b *satisfy conditions* (MA) *with* $p > d$. *For every* Q_i *fix a point* $x_0^i \in Q_i$. *Given* $t \in (t_n, t_{n+1})$ *and* $x \in Q_i$*, we define:*

$$\bar{\Phi}_E(t,x) := \bar{\Phi}_E(t_n,x) + (t - t_n) \fint_{t_n}^{t_{n+1}} \fint_{Q_i} b(s, \bar{\Phi}_E(t_n, x_0^i) + y - x_0^i)\,dy\,ds, \tag{5.4.1}$$

where $0 = t_0 < t_1 < \cdots < t_M = T$ *is a partition of* $[0,T]$ *with* $t_{i+1} - t_i = \Delta t$ *for some* $\Delta t > 0$ *and* $\Phi_E(0,x) = x$ $\forall x \in \mathbb{T}^d$. *In order to keep the notation as light as possible, we omit explicitly writing the dependence of* $\bar{\Phi}_E$ *on* Δt.

Before proving the stability estimate between Φ and $\bar{\Phi}_E$, we believe it is convenient to explicitly explain what the idea behind the definition of $\bar{\Phi}_E$ is and what approximation of u it can yield. Since the argument does not depend on the particular element of the partition, we drop the subscripts. Consider a point $x \in Q$. In the time interval $[0, \Delta t]$, such a point is transported by $\bar{\Phi}_E$ to

$$\bar{\Phi}_E(\Delta t, x) = x + \int_0^{\Delta t} \fint_Q b(s, y) dy ds = x + v.$$

This translation by v is applied to every point in Q, so in the first time step we are just translating Q by the vector $v = \int_0^{\Delta t} \fint_Q b(s, y) dy ds$. Let $\tilde{Q} := Q + v$. In the second time step, we have defined $\bar{\Phi}_E$ so that it is constant on $\tilde{Q}$ and it acts on $\tilde{Q}$ by translating it by $\int_{\Delta t}^{2\Delta t} \fint_{\tilde{Q}} b(s, y) dy ds$. Indeed, consider $\tilde{x} \in \tilde{Q}$. There exists $x \in Q$ such that $\tilde{x} = \bar{\Phi}_E(\Delta t, x) = x + v \in \tilde{Q}$. Then for every $\tilde{x} \in \tilde{Q}$:

$$\begin{aligned}
\bar{\Phi}_E(2\Delta t, x) &= \bar{\Phi}_E(\Delta t, x) + \int_{\Delta t}^{2\Delta t} \fint_Q b(s, \bar{\Phi}_E(\Delta t, x) + y - x) dy ds \\
&= \tilde{x} + \int_{\Delta t}^{2\Delta t} \fint_Q b(s, y + v) dy ds = \tilde{x} + \int_{\Delta t}^{2\Delta t} \fint_{\tilde{Q}} b(s, z) dz ds = \tilde{x} + w
\end{aligned}$$

with w not dependent on $\tilde{x}$. This is, in a sense, the same thing we have done when defining Φ_E. The difference is that while with Φ_E we were only integrating b in time (when $p > d$) and we kept the space variable fixed, here we also average over some translation of Q. To sum up, if we partition the initial datum as:

$$u_0 = \sum_{i=1}^N u_0 \llcorner Q_i = \sum_{i=1}^N u_0^i,$$

the proposed approximation of $u(t, \cdot)$ via $\bar{\Phi}_E$ will simply be:

$$[\bar{\Phi}_E(t, \cdot)]_\# u_0 = \sum_{i=1}^N [\bar{\Phi}_E(t, \cdot)]_\# u_0^i = \sum_{i=1}^N u_0^i(x - \bar{\Phi}_E(t, x_0^i)).$$

We now prove the stability estimate between Φ and $\bar{\Phi}_E$.

Proposition 5.4.1.
Consider $\{Q_i\}_{i=1}^N$ a partition of $\mathbb{T}^d$ such that $\mathrm{diam}(Q_i) \le \Delta x$ for every i, for some $\Delta x > 0$. Let b be a vector field satisfying (MA) *with $p > d$ and let $\bar{\Phi}_E$ be defined as in Definition 5.4.1 with time step Δt. It holds that:*

$$\|\Phi(t, \cdot) - \bar{\Phi}_E(t, \cdot)\|_{L^p} \lesssim C_t |\log(\max\{\Delta t, \Delta x\})|^{-1}$$

with $C_t \to 0$ as $t \to 0$.

Proof. Let $h := \max\{\Delta t, \Delta x\}$. We mimic the proof of Proposition 5.2.3 with $\bar{\Phi}_E$ in place of Φ_E. We recall that the following is a pointwise inequality, which holds for every $x, x_0 \in Q$ with $Q \in \{Q_i\}_{i=1}^N$ and for this reason we write $\Phi(t), \bar{\Phi}_E(t)$ in place of $\Phi(t, x), \bar{\Phi}_E(t, x)$. We have that:

$$\begin{aligned}
&\log\left(1 + \frac{|\Phi(t_n) - \bar{\Phi}_E(t_n)|}{h}\right) - \log\left(1 + \frac{|\Phi(t_{n-1}) - \bar{\Phi}_E(t_{n-1})|}{h}\right) \\
&\le \frac{h}{|\Phi(t_{n-1}) - \bar{\Phi}_E(t_{n-1})| + h} \frac{|\,|\Phi(t_n) - \bar{\Phi}_E(t_n)| - |\Phi(t_{n-1}) - \bar{\Phi}_E(t_{n-1})|\,|}{h} \\
&\le \frac{|\Phi(t_n) - \Phi(t_{n-1}) - (\bar{\Phi}_E(t_n) - \bar{\Phi}_E(t_{n-1}))|}{|\Phi(t_{n-1}) - \bar{\Phi}_E(t_{n-1})| + h} \\
&= \frac{\left|\int_{t_{n-1}}^{t_n} b(s, \Phi(s)) - \left(\fint_Q b(s, \bar{\Phi}_E(t_{n-1}, x_0) + y - x_0) dy\right) ds\right|}{|\Phi(t_{n-1}) - \bar{\Phi}_E(t_{n-1})| + h}.
\end{aligned}$$

For the sake of readability, we denote:

$$h_j(y) := \bar{\Phi}_E(t_j, x_0) + y - x_0.$$

By telescopic summing, as in (5.2.6), we end up having:

$$\log\left(1+\frac{|\Phi(t_n)-\bar{\Phi}_E(t_n))|}{h}\right) \le \sum_{j=0}^{n-1} \frac{\left|\int_{t_j}^{t_{j+1}} b(s,\Phi(s)) - \left(\fint_Q b(s,h_j(y))dy\right) ds\right|}{|\Phi(t_j)-\bar{\Phi}_E(t_j)|+h} \tag{5.4.2}$$

$$\le \sum_{j=0}^{n-1} \frac{|\int_{t_j}^{t_{j+1}} b(s,\Phi(s)) - b(s,\Phi(t_j))ds| + \left|\int_{t_j}^{t_{j+1}} b(s,\Phi(t_j)) - \left(\fint_Q b(s,h_j(y))dy\right) ds\right|}{|\Phi(t_j)-\bar{\Phi}_E(t_j)|+h}$$

$$\le \underbrace{\sum_{j=0}^{n-1} \frac{\int_{t_j}^{t_{j+1}} |b(s,\Phi(s)) - b(s,\Phi(t_j))|ds}{|\Phi(t_j)-\bar{\Phi}_E(t_j)|+h}}_{=1} + \underbrace{\sum_{j=0}^{n-1} \frac{\int_{t_j}^{t_{j+1}} \left|b(s,\Phi(t_j)) - \left(\fint_Q b(s,h_j(y))dy\right)\right| ds}{|\Phi(t_j)-\bar{\Phi}_E(t_j)|+h}}_{=2}.$$

The term 1 can be estimated exactly as in Proposition 5.2.3. As for 2, we use that $\mathrm{diam}(Q) \le \Delta x$, so arguing as in (5.2.8):

$$\frac{\left|b(s,\Phi(t_j)) - \fint_Q b(s,h_j(y))dy\right|}{|\Phi(t_j)-\bar{\Phi}_E(t_j)|+h} \le \frac{\fint_Q \left|b(s,\Phi(t_j)) - b(s,h_j(y))\right| dy}{|\Phi(t_j)-\bar{\Phi}_E(t_j)|+h} \tag{5.4.3}$$
$$\lesssim \frac{\fint_Q |\Phi(t_j)-h_j(y)|dy}{|\Phi(t_j)-\bar{\Phi}_E(t_j)|+h} f(s,\Phi(t_j)) \le f(s,\Phi(t_j)),$$

where we used the anisotropic estimate of Lemma 4.1.3, f is defined as in (5.2.9) and for the last inequality we used the fact that for every $y \in Q$:

$$|\Phi(t_j)-h_j(y)| \le |\Phi(t_j)-\bar{\Phi}_E(t_j)| + |y-x_0| \le |\Phi(t_j)-\bar{\Phi}_E(t_j)| + \Delta x. \tag{5.4.4}$$

We conclude that

$$\sup_{t\in[0,T]} \left\|\log\left(1+\frac{|\bar{\Phi}_E(t,\cdot)-\Phi(t,\cdot)|}{h}\right)\right\|_{L^p} \le C_t,$$

with $C_t \to 0$ as $t \to 0$ arguing as in Proposition 5.2.3, and given this bound on the L^p norm we can conclude as in Proposition 4.2.2. □

Remark 5.4.1. ***The restriction*** $p > d$
The anisotropic estimate of Lemma 4.1.3 is crucial for our proof of Proposition 5.4.1, and that is the reason why we only considered the case $p > d$. *If* $1 < p \le d$ *we do not have such an estimate, and we cannot estimate 2 using Lemma 4.1.2. Indeed, contrary to what we could do with* Φ_E, *we cannot change variables when using* $\bar{\Phi}_E$, *since it is clearly not injective. A more refined argument may solve the issue for* $1 < p \le d$, *but this will need further investigations.*

We can now prove the following.

Theorem 5.4.1. ***Diffuse deterministic approximation of*** u
Consider a partition $\{Q_i\}_{i=1}^N$ *of* $\mathbb{T}^d$, *an initial density* u_0 *and a vector field* b *such that* $\mathrm{diam}(Q_i) \le \Delta x$ *for every* i *and* b *satisfies* (MA) *with* $p > d$. *Then, given* $\bar{\Phi}_E$ *as in Definition 5.4.1 with a time step* Δt *and a function* $\bar{u}_0$ *such that* $\|\bar{u}_0 - u_0\|_{L^1(\mathbb{T}^d)} \le \varepsilon$ *for some* $\varepsilon > 0$, *it holds that*

$$\mathrm{W}_1(u(t,\cdot),[\bar{\Phi}_E(t,\cdot)]_\#\bar{u}_0) \lesssim \max\{C_t|\log(\max\{\Delta t,\Delta x\})|^{-1},\varepsilon\}.$$

Consider $\widetilde{\mathrm{W}}_1$ *the Wasserstein distance defined in Definition 5.3.1 with* $\alpha = 1$. *It holds that*

$$\widetilde{\mathrm{W}}_1(u(t,\cdot),(\bar{\Phi}_E(t,\cdot))_\#\bar{u}_0) \lesssim \max\{C_t,|\log(\max\{\Delta t,\Delta x\})|\varepsilon\}.$$

Proof. Let $h := \max\{\Delta t, \Delta x\}$. By the triangle inequality we have

$$\mathrm{W}_1(u(t,\cdot),[\bar{\Phi}_E(t,\cdot)]_\# \bar{u}_0) \leq \mathrm{W}_1(\overbrace{\Phi(t,\cdot)_\# u_0}^{=u(t,\cdot)}, \Phi(t,\cdot)_\# \bar{u}_0) + \mathrm{W}_1(\Phi(t,\cdot)_\# \bar{u}_0, [\bar{\Phi}_E(t,\cdot)]_\# \bar{u}_0). \tag{5.4.5}$$

We can estimate the first term with $\|\bar{u}_0 - u_0\|_{L^1} \leq \varepsilon$ and the second one with

$$\int_{\mathbb{T}^d} \bar{u}_0 |\Phi - \bar{\Phi}_E| dx \leq \|\bar{u}_0\|_{L^q(\mathbb{T}^d)} \|\Phi - \bar{\Phi}_E\|_{L^p(\mathbb{T}^d)} \lesssim C_t |\log(h)|^{-1}$$

by Proposition 5.4.1. As for the estimate for $\widetilde{\mathrm{W}}_1$, it follows in the exact same way. We just notice that 1-Lipschitz functions on $\mathbb{T}^d$ with respect to the distance d_1, defined in (5.3.3) with $\alpha = 1$, are uniformly bounded by $\approx |\log(h)|$ so we estimate the first addendum in (5.4.5) (with $\widetilde{\mathrm{W}}_1$ in place of W_1) with $|\log(h)|\varepsilon$, implicitly using Kantorovich-Rubinstein duality. In the same way we can estimate the second addendum with:

$$\begin{aligned}\widetilde{\mathrm{W}}_1(\Phi(t,\cdot)_\# \bar{u}_0, [\bar{\Phi}_E(t,\cdot)]_\# \bar{u}_0) &\leq \int_{\mathbb{T}^d} \bar{u}_0(x) \log\left(1 + \frac{|\Phi(t,x) - \bar{\Phi}_E(t,x)|}{h}\right) dx \\ &\leq \|\bar{u}_0\|_{L^q(\mathbb{T}^d)} \left\| \log\left(1 + \frac{|\Phi(t,\cdot) - \bar{\Phi}_E(t,\cdot)|}{h}\right)\right\|_{L^p(\mathbb{T}^d)} \lesssim C_t\end{aligned}$$

by Proposition 5.4.1, and the proof is complete. □

We remark that this method is very easily implemented numerically and parallelisable. Moreover, it provides a better rate of convergence for the logarithmic Wasserstein distance than the one proved in [142] (assuming $\Delta t = \Delta x = h/2$ for some $h > 0$), having h in the denominator in the definition of the logarithmic Wasserstein distance rather than $\sqrt{h}$. We finally remark that such an improvement in the case $p > d$ is in accordance with what we proved in Theorem 5.3.3.

Remark 5.4.2. ***The lack of a CFL condition and a time step limitation for unbounded vector fields***
A usual drawback of explicit methods is the so-called CFL condition, i.e. the requirement that $\|b\|_\infty \Delta t/\Delta x \lesssim 1$. In this case, contrary to what we remarked in [56], there is no such condition. In section 5.3 we worked essentially in a meshless setting, and the partition $\{Q_i\}_{i=1}^N$ had the only role of providing an approximation of the initial datum by selecting a Dirac mass with suitable weight inside each Q_i. We had either a situation where both the time step Δt and the space discretization scale Δx contributed to the total error separately (Theorem 5.3.2) or a situation where we gave estimates with respect to the maximum of the two (Theorem 5.3.3). Similar estimates also apply to the diffuse case, where the error depends on $h = \max\{\Delta t, \Delta x\}$. We remark that the lack of a CFL condition may derive from the different point of view employed: while Eulerian methods (like the finite volume method) seem to be inherently affected by the need of a CFL condition, Lagrangian methods may have the advantage of not needing this additional requirement. We refer the reader to [94] for a broader discussion on such a matter (in the case of regular vector fields).

Another issue is the stability of the method at the level of flows. We proved that no condition on Δt is needed when b is bounded. However, from numerical experiments (and classical numerical analysis theory) it is clear that the situation is different when b is not bounded. Namely, a time step too big may result in instabilities which make the error blow up. We do not have an explicit proof in the non-regular case or bounds on Δt, since this would require stability results at the level of approximate flows similar to the ones obtained in [58, Section 3] where the authors drop the boundedness assumption. We believe that such estimates are possible, with some additional work, mimicking the proofs of the bounded case.

Remark 5.4.3. ***Weak vs. strong approximations of*** u_0
In Section 5.3 we worked with a sum of Dirac deltas which approximated, in a weak sense, the initial datum u_0. Choosing such a measure was very natural, since we wanted a "true" approximation method, in the sense that it is truly implementable in a computer. Indeed in sections 5.2 and 5.3 we used Φ_E which is not locally constant in the space variable, so the only possibility was to use measures concentrated on points.

In Section 5.4, working with $\bar{\Phi}_E$ which is locally constant in space, we are more free to choose the initial datum, and actually we can compute the push-forward of every initial datum $\bar{u}_0$ as we explained after Definition 5.4.1. In Theorem 5.4.1 we imposed that $\bar{u}_0$ approximated u_0 in a strong L^1 sense, so it is natural to wonder if the same result holds only assuming $\bar{u}_0$ to be a measure approximating u_0 in a weak sense (for instance, one could consider $\bar{u}_0$ to be a sum of Dirac deltas just like in section 5.3 and hope to obtain a deterministic result). Given W the Wasserstein distance with respect to some distance d, in order to estimate $\mathrm{W}(u(t,\cdot), [\bar{\Phi}_E(t,\cdot)]_{\#}\bar{u}_0) = \mathrm{W}(\Phi(t,\cdot)_{\#}u_0, [\bar{\Phi}_E(t,\cdot)]_{\#}\bar{u}_0)$ we crucially rely on the triangle inequality. Namely, we need to estimate either $\mathrm{W}(\Phi(t,\cdot)_{\#}u_0, \Phi(t,\cdot)_{\#}\bar{u}_0)$ or $\mathrm{W}([\bar{\Phi}_E(t,\cdot)]_{\#}u_0, [\bar{\Phi}_E(t,\cdot)]_{\#}\bar{u}_0)$. The problem is that, to the best of our knowledge, in general such distances can only be estimated with the L^1 norm of the difference between u_0 and $\bar{u}_0$. If f is a Lipschitz function it actually holds that $\mathrm{W}(f_{\#}\mu, f_{\#}\nu) \le \mathrm{Lip}(f)\,\mathrm{W}(\mu,\nu)$, however neither Φ nor $\bar{\Phi}_E$ are Lipschitz. And even if we previously approximated b with some Lipschitz function $\tilde{b}$, the Lipschitz constant of the flow of $\tilde{b}$ would depend exponentially on both $\mathrm{Lip}(\tilde{b})$ and the time t, which is clearly not ideal.

5.5 Approximation via θ-method and numerical experiments

In this section we give the main results proved in [46], which consist in approximation results for the regular Lagrangian flow via a θ-method. With such estimates on the flow, as we have already seen, we can recover stability estimates for solutions of (PDE). First of all, since it is less-known than the forward Euler method, let us define the classical θ-method.

Definition 5.5.1. ***The classical θ-method***
Let $T > 0$ and let $b : [0,T) \times \mathbb{R}^d_x \to \mathbb{R}^d_x$ be a smooth and bounded vector field. Denote with $X(t,x)$ the flow of b starting at x when $t = 0$, i.e. let $X(\cdot,x)$ solve:

$$\begin{cases} \frac{d}{dt}X(t,x) = b(t,X(t,x)) & \text{if } t > 0 \\ X(0,x) = x & \text{if } t = 0 \end{cases}.$$

Given a parameter $\theta \in [0,1]$ and a time step $\Delta t = T/N$ for some $N \in \mathbb{N}$, we consider the net $\{t_i\}_{i=0}^N$ defined by $t_i := i\Delta t$ for $i = 0,\dots,N$. The classical θ-method, which produces a sequence $\{X_i^\theta\}_{i=0}^N$ approximating $X(t_i,x)$ for every $i = 0,\dots,N$ is given by (we omit explicitly writing the dependence on x):

- *$X_0^\theta := x$;*
- *for every $i = 0,\dots,N-1$, given X_i^θ we find X_{i+1}^θ by solving*

$$\frac{X_{i+1}^\theta - X_i^\theta}{\Delta t} = (1-\theta)b(t_i, X_i^\theta) + \theta b(t_{i+1}, X_{i+1}^\theta). \tag{5.5.1}$$

In particular, if $\theta = 0$ or $\theta = 1$ the θ-method coincides with the forward Euler method and the backward Euler method, respectively [3].

Just like the forward Euler method, if b is not regular neither in space nor in time the definition in (5.5.1) does not make sense. So, in the setting of vector fields with $L^1_t W^{1,p}_x$ regularity (with $p > 1$), we will use an approximation of b in space and a time average to define the iterative steps of the method. Namely, we consider an approximating sequence $\{b_\varepsilon\}_{\varepsilon>0}$ satisfying the following properties [4]:

[3] If $\theta = 1/2$ the method coincides with the *Crank-Nicolson* method, which for C^2 vector fields gives an error of order $O(\Delta t^2)$, while for every other value of θ the method is only first-order accurate.

[4] The standard regularisation via mollification, in particular, satisfies these assumptions. Another example of an approximating sequence is provided in [46, Appendix A].

$$\|b_\varepsilon - b\|_{L^1_t L^1_x} \le \bar{C}_0 \varepsilon^\alpha, \tag{5.5.2}$$
$$\|\nabla b_\varepsilon\|_{L^\infty_t L^\infty_x} \le \frac{\bar{C}_1}{\varepsilon^\beta}, \tag{5.5.3}$$
$$\|\operatorname{div}_x b_\varepsilon\|_{L^\infty_t L^\infty_x} \le \bar{C}_2, \tag{5.5.4}$$
$$\|b_\varepsilon\|_{L^\infty_t L^\infty_x} \le \bar{C}_3, \tag{5.5.5}$$

for some $\bar{C}_0, \bar{C}_1, \bar{C}_2, \bar{C}_3, \alpha, \beta > 0$.

We can now define the *regularised* θ-method and state the main results obtained in [46]. We skip the proofs since they are similar to the ones used for the forward Euler method in previous sections.

Definition 5.5.2. ***Regularised θ-method***
For $x \in \mathbb{R}^d$, set $X_0^{\theta,\varepsilon}(x) = x$, and for $i = 0, \dots, N-1$, given $X_i^{\theta,\varepsilon} = X_i^{\theta,\varepsilon}(x)$, compute $X_{i+1}^{\theta,\varepsilon} = X_{i+1}^{\theta,\varepsilon}(x)$ (as previously, we omit explicitly writing the dependence on x when not necessary) by solving

$$\frac{X_{i+1}^{\theta,\varepsilon} - X_i^{\theta,\varepsilon}}{h} = (1-\theta) \fint_{t_i}^{t_{i+1}} b_\varepsilon(s, X_i^{\theta,\varepsilon}) ds + \theta \fint_{t_i}^{t_{i+1}} b_\varepsilon(s, X_{i+1}^{\theta,\varepsilon}) ds. \tag{5.5.6}$$

We define $\tilde{X}^{h,\theta,\varepsilon}$ by interpolating the sequence $\{X_i^{\theta,\varepsilon}\}_{i=0}^N$:

$$\tilde{X}^{h,\theta,\varepsilon}(t) := \sum_{i=0}^{N-1} \left(X_i^{\theta,\varepsilon} + \frac{t-t_i}{h}(X_{i+1}^{\theta,\varepsilon} - X_i^{\theta,\varepsilon}) \right) \chi_{[t_i,t_{i+1})}(t). \tag{5.5.7}$$

Remark 5.5.1.
In order to prove that a solution to the regularised θ-method exists when $\theta > 0$ (i.e. when the method is not completely explicit), we need to impose that $h\|\nabla b_\varepsilon\|_\infty \le 1/2$ (see [46, Proposition 3.1]), although it is not clear to us if this assumption is needed or if it is just a shortcoming of our proof.

The two main results obtained in [46] are given next:

Theorem 5.5.1. ***Stability estimate for the θ-method, [46]***
Let $b \in L^1((0,T); W^{1,p}(\mathbb{R}^d)) \cap L^\infty((0,T)\times\mathbb{R}^d)$ be a divergence-free velocity field for some $1 < p < \infty$. For $\varepsilon > 0$, let $\{b_\varepsilon\}_{\varepsilon>0}$ be an approximating sequence of the velocity field b satisfying (5.5.2)-(5.5.5) *and let $h := \max\{\varepsilon^\alpha, \varepsilon^{2\beta}\}$. Then, for any $r > 0$, there exists a positive constant C such that, given $\tilde{X}^{h,\theta,\varepsilon}$ defined as in* (5.5.7)*:*

$$\|\tilde{X}^{h,\theta,\varepsilon}(t,x) - X(t,x)\|_{L^1(B_r)} \le \frac{C}{|\log h|}. \tag{5.5.8}$$

In particular, C depends on $d, p, r, T, \|b\|_{L^\infty}, \|\nabla b\|_{L^1_t L^p_x}, \bar{C}_1, \bar{C}_2$.

Before stating the next result we need to introduce an approximating sequence for the backward regular Lagrangian flow of b (see Definition 2.2.5). For every $t \in [0,T]$ we consider the backward regular Lagrangian flow Y_t solving

$$\begin{cases} \partial_s Y_t(s,x) = -b(t-s, Y_t(s,x)) \\ Y_t(0,x) = x \end{cases} \tag{5.5.9}$$

so that, if X is the regular Lagrangian flow of b, we have the relation $X^{-1}(t,\cdot) = Y_t(t,\cdot)$. With these notations, (5.5.9) is a forward ODE which keeps track of the final time t. Then, we can write the θ-method for (5.5.9) as follows. Let $\theta \in [0,1]$, and for any fixed $N \in \mathbb{N}$ define the time step $h := T/N$ and consider $t = jh$ for some $j = 0, \dots, N$. Consider the net $\{t_i\}_{i=0}^j$ defined by $t_i := ih$ for $i = 0, \dots, j$. For simplicity, we also define

$$\bar{b}_t(s,x) = \begin{cases} -b(t-s,x) & \text{if } s \le t \\ 0 & \text{otherwise} \end{cases}. \tag{5.5.10}$$

Then, we proceed iteratively as follows. Set $Y_0^{\theta,\varepsilon}(t,x) = x$. For $i \in \{0,\dots,j-1\}$, given $Y_i^{\theta,\varepsilon} = Y_i^{\theta,\varepsilon}(t,x)$, we find $Y_{i+1}^{\theta,\varepsilon} = Y_{i+1}^{\theta,\varepsilon}(t,x)$ by solving

$$\frac{Y_{i+1}^{\theta,\varepsilon} - Y_i^{\theta,\varepsilon}}{h} = (1-\theta) \fint_{t_i}^{t_{i+1}} \bar{b}_t^\varepsilon(\tau, Y_i^{\theta,\varepsilon}) d\tau + \theta \fint_{t_i}^{t_{i+1}} \bar{b}_t^\varepsilon(\tau, Y_{i+1}^{\theta,\varepsilon}) d\tau, \tag{5.5.11}$$

where we omit writing explicitly the dependence on t and x and where $\bar{b}_t^\varepsilon$ is an approximating sequence for $\bar{b}_t$ satisfying (5.5.2)-(5.5.5). Notice that if b_ε is an approximating sequence for b, then $\bar{b}_t^\varepsilon$ is an an approximating sequence for $\bar{b}_t$, where $\bar{b}_t^\varepsilon$ is defined in the obvious way. Given the sequence $\{Y_i^{\theta,\varepsilon}\}_{i=0}^{j-1}$ constructed in (5.5.11), we define the function

$$\tilde{Y}^{N,\theta,\varepsilon}(t,s,x) := \sum_{i=0}^{N-1} \left(Y_i^{\theta,\varepsilon} + \frac{s-t_i}{h}(Y_{i+1}^{\theta,\varepsilon} - Y_i^{\theta,\varepsilon}) \right) \chi_i^t(s), \tag{5.5.12}$$

where with χ_i^t we denote the characteristic function of the set $[0,t] \cap [t_i, t_{i+1}[$,. Finally, we define the sequence u^h of approximating solutions of (PDE) (which by the divergence-free assumption on b is equivalent to a transport equation) as follows:

$$u^h(t,\cdot) := u_0(\tilde{Y}^{N,\theta,\varepsilon}(t,t,\cdot)), \tag{5.5.13}$$

where to simplify the notations we drop the superscripts θ, ε on the function u^h. The following holds:

Theorem 5.5.2. ***Error estimates for the Lagrangian solution to the transport equation***
Let $b \in L^1((0,T); W^{1,p}(\mathbb{R}^d)) \cap L^\infty((0,T)\times\mathbb{R}^d)$ be a divergence-free velocity field for some $1 < p < \infty$ and let $\{Y_t\}_{t\in[0,T)}$ be the backward regular Lagrangian flows solving (5.5.9)*, for every $t \in [0,T)$. Define the sequence $\{Y_i^{\theta,\varepsilon}\}_{i,\varepsilon}$ as in* (5.5.11) *and $\tilde{Y}^{h,\theta,\varepsilon}$ as in* (5.5.12)*, corresponding to an approximating sequence $\bar{b}_t^\varepsilon$ of the* backward vector field $\bar{b}_t$, *as in* (5.5.10). *Then, there exists a constant $C = C(d,p,T,\|b\|_{L^\infty_{t,x}}, \|\nabla b\|_{L^1_t L^p_x}, \bar{C}_1, \bar{C}_2)$ such that for any $u_0 \in L^1(\mathbb{R}^d)$ the sequence u^h defined in* (5.5.13)*, constructed via the backward flow $\tilde{Y}^{h,\theta,\varepsilon}$, converges to the unique Lagrangian solution* [5] *u^L of* (PDE)*, with the estimate*

$$\|u^h(t,\cdot) - u^L(t,\cdot)\|_{L^1} \le C \left(\|u_0^\delta - u_0\|_{L^1} + \frac{C_\delta}{|\log h|} \right), \tag{5.5.14}$$

where $h := \max\{\varepsilon^\alpha, \varepsilon^{2\beta}\}$, u_0^δ is any smooth approximation of u_0 with compact support and Lipschitz constant δ^{-1} and the constant C_δ blows up as $\delta \to 0$.

Remark 5.5.2.
We remark that the incompressibility assumption was crucial to get strong L^1 estimates in the previous theorem. If b is not divergence-free, assumptions on $D_x \operatorname{div}_x b$ seem necessary (see Proposition 4.3.2).

Let us conclude this section with selected numerical experiments. To this purpose, we consider the velocity field

$$b(x_1,x_2) = \left(-2(\alpha+1)(x_1^2+x_2^2)^{\frac{\alpha-1}{2}} x_2,\ 2(\alpha+1)(x_1^2+x_2^2)^{\frac{\alpha-1}{2}} x_1 \right) \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^2), \tag{5.5.15}$$

where $1 - \frac{2}{p} < \alpha < 1$ and $p > 2$. The vector field in (5.5.15) is smooth away from the origin, where it has a singularity, so no spatial regularisation is needed. The associated ODE system admits the explicit solution

$$\begin{cases} x_1(t) = r_0 \cos\left(\beta_0 + 2(\alpha+1) r_0^{\alpha-1} t\right) \\ x_2(t) = r_0 \sin\left(\beta_0 + 2(\alpha+1) r_0^{\alpha-1} t\right) \end{cases} \tag{5.5.16}$$

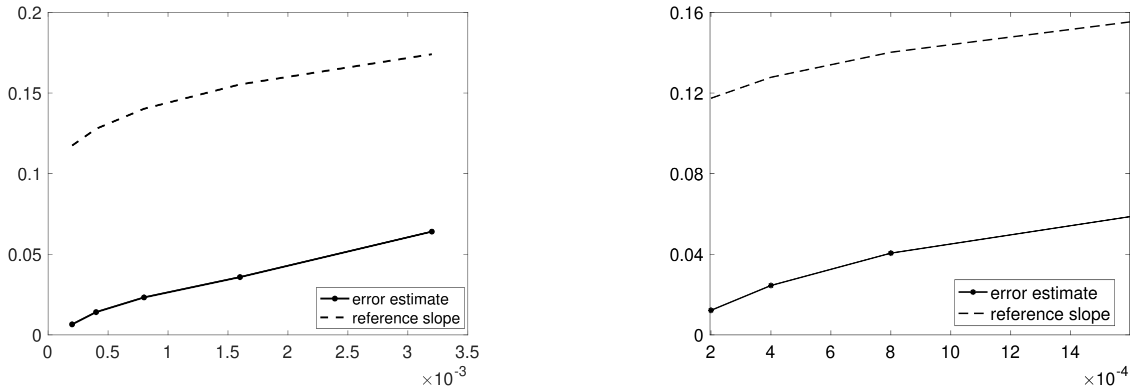
where $r_0 = \sqrt{x_1(0)^2 + x_2(0)^2}$ and $\beta_0 = \arctan\left(\frac{x_2(0)}{x_1(0)}\right)$. We apply the θ-method (5.5.6) for selected values of h and $\theta \in [0,1]$. Depending on the choice of the initial condition, we observe the following behavior:

[5] It is not necessary for u_0 to satisfy the integrability requirements needed by the DiPerna-Lions to have uniqueness of solutions. Even outside of the DiPerna-Lions range we have convergence to the unique Lagrangian solution (see Definition 2.2.2), which may not be the unique weak solution of (PDE).

- When the initial condition is placed sufficiently far from the origin, the convergence rate appears to be faster than $1/|\log h|$, yet still slower than linear. Moreover, the trajectories obtained by the θ-method are circles, in agreement with the exact solution of the system.

- When the initial datum is chosen in a small neighborhood of the origin, the observed convergence rate aligns with the logarithmic rate predicted by our theoretical analysis.

Figure 5.1 displays the pattern of the error estimate and highlights its decay in comparison with a reference logarithmic slope.

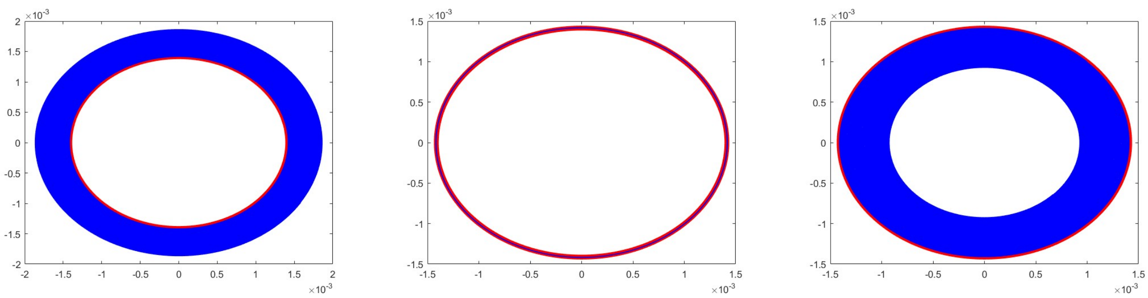


(a) $\theta = 0.2$, $p = 3$, $\alpha = 0.36$.

(b) $\theta = 0$, $p = 3$, $\alpha = 0.35$.

Figure 5.1: Error estimate computed according to (5.5.8), arising from the application of the θ-method to the velocity in (5.5.15) with initial condition $(0.001, 0.001)$ and a final time $T = 4$. The reference solution is (5.5.16) and the displayed error decay results from the application of the θ-method with step sizes $h = 2^k h_0$ for $k = 0, 1, 2, 3$, with $h_0 = 10^{-4}$.

In the plots, we compute the difference between the trajectory (with step size h) and the reference trajectory given by (5.5.16). This difference is evaluated assuming a fixed initial position $(0.001, 0.001)$, and should be interpreted as a pointwise error between trajectories, rather than a norm-based error (such as a L^1 norm). Nevertheless, this pointwise comparison is equivalent to evaluating the error in the $L^\infty_t L^1_x$ norm, since when implementing a quadrature rule to approximate a spatial integral (e.g., the L^1 norm) we evaluate the difference between the two solutions at mesh points. Thus, the pointwise distance between trajectories at discrete times naturally corresponds to a discrete approximation of the desired norm. Concerning the numerical trajectories, they exhibit a spiraling behavior: they move away from the origin when $\theta < 1/2$, and converge towards the origin when $\theta > 1/2$. Instead, for $\theta = 1/2$, the trajectories are circles, see Figure 5.2.



(a) $\theta = 0.49$, $p = 3$, $\alpha = 0.34$.

(b) $\theta = 0.5$, $p = 3$, $\alpha = 0.34$.

(c) $\theta = 0.51$, $p = 3$, $\alpha = 0.34$.

Figure 5.2: Approximate trajectories (in blue) for the vector field (5.5.15) for different values of θ with step size $h = 10^{-4}$, initial datum $(0.001, 0.001)$ and integrated for a time $T = 4$, versus the exact solution (5.5.16) (in red, with a thicker line). The approximate trajectories are very dense, so the individual lines cannot be distinguished.

We also consider the vector fields

$$b(x_1, x_2) = \left(\sqrt{|\sin 2\pi x_2|},\ \sqrt{|\sin 2\pi x_1|}\right) \in W^{1,q}_{\mathrm{loc}}(\mathbb{R}^2) \text{ with } 1 \le q < 2, \tag{5.5.17}$$

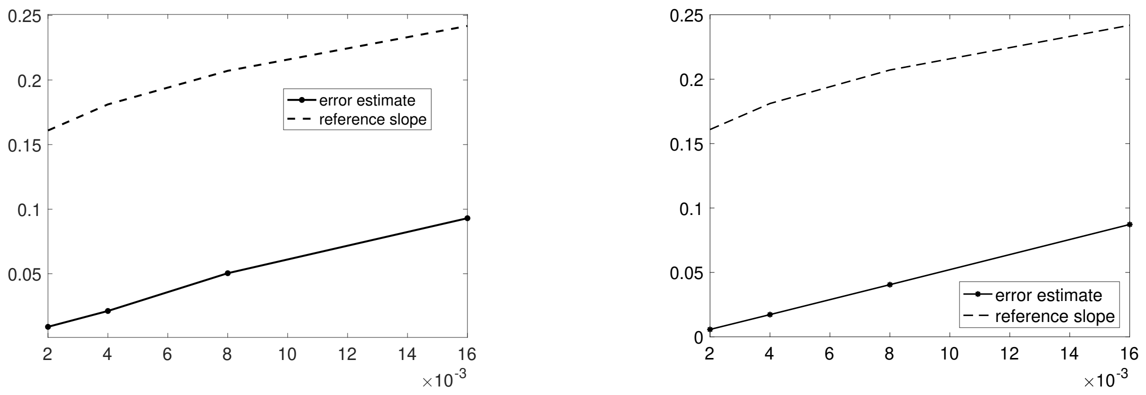
and

$$b(x_1, x_2) = \left(\log\left(\frac{1}{|x_2|}\right)^{\frac{1}{p}} |x_2|^{1-\frac{1}{2p}}, \log\left(\frac{1}{|x_1|}\right)^{\frac{1}{p}} |x_1|^{1-\frac{1}{2p}}\right) \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^2), \tag{5.5.18}$$

for some $p > 1$ to be chosen, and initial datum in a small neighbourhood of the origin. Our numerical procedure is as follows.

1. We fix a minimum step size h_0 and solve the ODE using the θ-method.
2. We increase the step size and treat the numerical solution corresponding to the smallest step size h_0 as our reference (exact) solution.

Figure 5.3 displays the pattern of the error estimate and highlights its decay in comparison with a reference logarithmic slope. In the plots, we consider the error of the θ-method to be given by the difference between the trajectory obtained with step size h and the reference trajectory corresponding to h_0. The logarithmic slope confirms the feasibility of the error estimate of Theorem 5.5.1.



(a) $\theta = 0.8$, $h_0 = 10^{-3}$.

(b) $\theta = 0.7$, $p = 100$, $h_0 = 10^{-3}$.

Figure 5.3: Error estimate computed according to (5.5.8), arising from the application of the θ-method to the vector fields (5.5.17) (on the left) and (5.5.18) (on the right). The reference solution is computed with step size h_0 and the displayed error decay results from the application of the same method with step sizes $h = 2^k h_0$, with $k = 0, 1, 2, 3$.

5.6 Final discussion, questions and future perspectives

We conclude by briefly summarising the results obtained, pointing out some questions and future possible research directions. In this chapter, we proved a stability estimate between the regular Lagrangian flow of a vector field b with Sobolev regularity in space and an approximation of such flow via a (sort of) explicit Euler method. We gave two approximations of the flow, i.e. Φ_E and $\bar{\Phi}_E$, for which the same stability estimate holds. We used that the solution u of the continuity equation (PDE) is representable as a push-forward of the initial datum u_0 via the regular Lagrangian flow Φ of b to approximate it using our approximations Φ_E and $\bar{\Phi}_E$ of Φ and different approximations of the initial datum u_0. In sections 5.2 and 5.3 we used Φ_E to approximate Φ and a sum of Dirac deltas to approximate u_0. In this setting, we proved that the Dirac deltas, advected by Φ_E, approximate u in suitable Wasserstein distances. However such an approximation holds only in expected value, since a priori it strongly depends on the choice of the Dirac deltas.

Question 1. *Is the estimate truly valid in expected value only, or it can be made deterministic if $h = \Delta t = \Delta x$ is small enough?*

In section 5.4, we used a different approximation $\bar{\Phi}_E$ of Φ and a strong approximation $\bar{u}_0$ of u_0 in L^1 norm to prove a deterministic estimate. However, due to technical reasons in our proof of Proposition 5.4.1, we can prove that the crucial estimate on $\|\Phi(t,\cdot) - \bar{\Phi}_E(t,\cdot)\|_{L^p(\mathbb{T}^d)}$ holds only assuming $p > d$.

Question 2. *Does the stability estimate on $\|\Phi(t,\cdot) - \bar{\Phi}_E(t,\cdot)\|_{L^p(\mathbb{T}^d)}$ also hold for $1 < p \leq d$?*

Moreover, as already said in Remark 5.4.3, we do not know if a strong approximation of u_0 is truly needed in Theorem 5.4.1:

Question 3. *Does Theorem 5.4.1 hold even assuming to have as initial datum a measure $\bar{u}_0$ approximating u_0 in a weak sense (for example in some Wasserstein distance)?*

Of course, it is totally possible that a better *explicit* approximation of Φ, that has nothing to do with either Φ_E or $\bar{\Phi}_E$ and for which there are no issues like the ones we just described above, exists. For example, in [46] we are considering an approximation of Φ based on a θ-method, which for $\theta > 0$ should have better stability properties even for unbounded vector fields.

To conclude, we highlight possible future research directions and recap the results we obtained. While numerical approximation methods for the continuity equation [19, 31, 96, 142, 143, 159] and more recently for the advection-diffusion equation [128] have received some attention, more general situations where there is a non-linear coupling between the density and the velocity field seem less studied. In this direction we cite [96], whose results can be applied also to some non-linear systems. However, a limitation of the available methods is a certain "rigidity" with respect to the mesh which seems somehow unavoidable (see [96, Example 1.1]) when considering an *explicit* finite volume scheme (we recall [143], where a Cartesian mesh is not needed, but using an *implicit* method). For this reason the mesh needs to be either Cartesian (as in [142]) or at least to satisfy some additional conditions (e.g. having a periodic structure [96]), which may be problematic in some situations. Our method does not need any assumption on the mesh other than an upper bound on the diameter of the cells and, eventually, a lower bound on their volume. Moreover it is easily parallelisable and it does not need to satisfy a CFL condition (at least if b is bounded), but it only provides weak stability estimates in Wasserstein distance.

Future challenges may be the approximation of the solution of coupled systems without relying on any particular structure of the mesh, or developing methods based on a Lagrangian point of view that approximate the solution u of (PDE) in a stronger topology.

Chapter 6

Advection-diffusion equation

6.1 Setting and main assumptions

In this chapter we want to show how the Lagrangian point of view can help also in dealing with the advection-diffusion equation:

$$\begin{cases} \partial_t u + b \cdot \nabla_x u = \nu \Delta_x u & \text{in } (0,T) \times \mathbb{T}^d \\ u(0,x) = u_0(x) \end{cases} . \tag{ADE}$$

For simplicity, we will work in the d-dimensional torus $\mathbb{T}^d$ and we will consider a vector field b bounded and incompressible. Moreover, we the assume the diffusivity $\nu > 0$ constant, which will be a crucial simplification (see Remark 6.3.2).

Before discussing (ADE), let us introduce some notation we will use in this chapter. We will work with a probability space $(\Omega, \mathcal{F}, \mathrm{P}_\omega)$, where $\mathcal{F}$ is a σ-algebra and P_ω is a probability on $\mathcal{F}$. We will use the notations $\mathbb{E}_\omega[\cdot]$ and $\mathrm{Var}_\omega[\cdot]$ to denote the expected value and the variance of random variables with respect to the probability P_ω. We will also use the notations $\mathbb{E}_{x_0}[\cdot]$ and $\mathrm{Var}_{x_0}[\cdot]$ to denote the expected value and the variance with respect to a probability $\bar{\mathrm{P}}_{x_0}$ that we will introduce in due time. We will sometimes make use of basic results on stochastic processes and probability, like Itô's lemma, Doob's maximal principle or standard properties of the Brownian motion. Such results are can be found in every book on SDEs, see for instance [15, 98, 131].

The interest on (ADE) (also referred to as Fokker-Planck [1] equation or convection-diffusion equation) stems from different fields, most notably fluid dynamics and finance, with (ADE) being related to Navier-Stokes equations and Black-Scholes equation. The study of (ADE) in presence of smooth vector fields dates back to the classical theory developed in [105], but in practice one often has to work with lower regularity assumptions, especially in physics where the velocity b may exhibit a turbulent behaviour [52, 147], so a mathematical theory for such settings is required. This was first addressed, to the best of our knowledge, in [109], and later in [77, 110]. The key observation is that (ADE) is strongly connected to the following (backward) stochastic differential equation:

$$\begin{cases} dX^\nu_{s,t}(x,\omega) = -b(t-s, X^\nu_t(x,\omega))ds + \sqrt{2\nu}dW_s(\omega), & s \in [0,t] \\ X^\nu_{t,t}(x,\omega) = x \end{cases} . \tag{B-SDE}$$

Notice that, if $\nu = 0$, this is exactly the backward flow introduced in Definition 1.1.2 for the classical case and in Definition 2.2.5 for the non-classical case. The link comes from the Feynman-Kac formula [103], which gives a representation of the solution u of (ADE) in terms of the solution $X^\nu_{s,t}$ of (B-SDE) and of the initial datum u_0 [2].

[1] Actually, the Fokker-Planck equation has $\mathrm{div}_x(bu)$ in place of $b \cdot \nabla_x u$, but since we are working with an incompressible vector field the two expressions are equivalent. In most of the cited papers, the equation has $\mathrm{div}_x(bu)$ in place of $b \cdot \nabla_x u$, not having the divergence-free assumption.

[2] We will not enter into details, but in the non-classical setting the Feynman-Kac representation of a solution holds for the so-called *parabolic solutions*, which are the same as the Lagrangian solutions for the continuity and transport equations. We will always have assumptions that guarantee that there exists a unique weak solution of (ADE) and it is parabolic. See [26] for more

If we consider the forward formulation of (B-SDE), i.e:

$$\begin{cases} dX_t^\nu(x,\omega) = b(t, X_t^\nu(x,\omega))dt + \sqrt{2\nu}dW_t(\omega), & t \in (0,T) \\ X_0^\nu(x,\omega) = x \end{cases}, \tag{SDE}$$

we can also use a push-forward representation in the Feynman-Kac formula:

$$u(t,x) = \mathbb{E}_\omega[u_0(X_{0,t}^\nu(x,\omega))] = \mathbb{E}_\omega[X_t^\nu(\cdot,\omega)_\# u_0]. \tag{F-K}$$

We refer the reader to Remark 2.3.1 for a discussion on the well-posedness of (SDE) when b is not regular, see also [140, 120, 121]. Up to our knowledge, the first quantitative stability estimates for solutions of (ADE) with non-regular velocity fields, in logarithmic Wasserstein distance, were provided in [115]. The authors use techniques inspired by [58], the representation (F-K) and the regularising properties of the Brownian motion not only to consider vector fields in the DiPerna-Lions setting, but also vector fields satisfying the Ladyzhenskaya-Prodi-Serrin condition, asking less regularity in the spatial variables but more integrability in the time, namely

$$b \in [L^q(0,T; L^p(\mathbb{T}^d))]^d \quad \text{with } p, q > 2 \text{ such that } \frac{d}{p} + \frac{2}{q} \leq 1.$$

We refer the reader to [115, Theorem 2.2, Theorem 2.6] for the precise statements. An improvement on the rate of convergence (still in logarithmic Wasserstein distance) in the DiPerna-Lions setting with $p > 1$ (i.e. $b \in L_t^1 W_x^{1,p}$ with $p > 1$ and $\mathrm{div}_x(b)^- \in L_t^1 L_x^\infty$) has been given in [129], where the representation formula (F-K) is not used and the argument completely relies on deterministic methods. In particular, if u_1 and u_2 are solutions of (ADE) relative to different diffusivities ν_1 and ν_2, in [115] the authors give an estimate depending on $\sqrt{|\nu_1 - \nu_2|}$, while the estimates in [129] depend on $|\nu_1 - \nu_2|$. For more details on well-posedness results for (ADE) in a non-regular setting we refer the reader to [111, 26].

We want to show how we can use Lagrangian estimates on the stochastic flow which solves (SDE) and the representation provided by (F-K) to derive estimates for the solution of (ADE), in a sense using the same point of view of [56, 46]. The main assumptions we will have on b are the following:

$$\begin{cases} b \in [L^1(0,T; W^{1,p}(\mathbb{T}^d)) \cap L^\infty((0,T)\times\mathbb{T}^d)]^d & \text{with } p > 1 \\ b \text{ is divergence-free} \end{cases}. \tag{6.1.1}$$

Remark 6.1.1. ***A selection principle for the Lagrangian solution of the transport equation***
We remark that, as $\nu \to 0$ and under the assumptions (6.1.1)*, the solution of* (ADE) *converges to the unique Lagrangian solution of the corresponding transport equation (i.e. with $\nu = 0$) even when the integrability of the initial datum is not sufficient to be in the DiPerna-Lions range where we have uniqueness of solutions. See [25] for more details.*

We can now define the *stochastic flow* [3]:

Definition 6.1.1. ***Stochastic flow***
Let $\nu > 0$ and let $b \in [L^\infty((0,T)\times\mathbb{T}^d)]^d$ be a divergence-free vector field. A map $X^\nu : [0,T)\times\mathbb{T}^d\times\Omega \to \mathbb{T}^d$ is a stochastic flow for b with diffusivity ν if:

- *for any $t \in [0,T]$, for any $x \in \mathbb{T}^d$ and for a.e. $\omega \in \Omega$, the map $t \mapsto X^\nu(t,x,\omega) = X_t^\nu(x,\omega) \in \mathbb{T}^d$ is an absolutely continuous solution of*

$$\begin{cases} dX_t^\nu(x,\omega) = b(t, X_t^\nu(x,\omega))dt + \sqrt{2\nu}dW_t(\omega) & \text{with } t \in (0,T) \\ X_t^\nu(x,\omega) = x \end{cases};$$

- *for every $t \in [0,T)$ and for a.e. $\omega \in \Omega$, the map $x \in \mathbb{T}^d \mapsto X_t^\nu(x,\omega)$ is measure preserving.*

For the proof of the existence an uniqueness of the stochastic flow under the assumptions (6.1.1) and its measure-preserving property we refer the reader to [45] and [78, Lemma 9].

details.

[3] Contrary to [25] we will only focus on weak estimates between solutions, whereas they considered a backward flow in order to use (F-K) to derive strong L^1 estimates.

6.2 Stability estimates for stochastic flows

We first give a result that is the analogous of Proposition 4.2.2 for stochastic flows and which is obtained essentially following the lines of [25, Lemma 5.3]:

Proposition 6.2.1. ***Stability estimate for the stochastic flow***
Consider X_t^1 and X_t^2 two stochastic flows with respect to velocity fields b_1, b_2 satisfying (6.1.1) *and with diffusivities ν_1, ν_2. Then:*

$$\int_{\mathbb{T}^d} \mathbb{E}_\omega[|X_t^1(x,\omega) - X_t^2(x,\omega)|]dx \lesssim C_t \left(\sqrt[4]{|\nu_1 - \nu_2|} + \sqrt{\|b_1 - b_2\|_{L^1(0,t;L^p(\mathbb{T}^d))}} \right.$$
$$\left. + \frac{1}{\left|\log\left(\sqrt{|\nu_1 - \nu_2|} + \|b_1 - b_2\|_{L^1(0,t;L^p(\mathbb{T}^d))}\right)\right|} \right).$$

Proof. For any $t \in (0,T)$, a.e. $\omega \in \Omega$ and a.e. $x \in \mathbb{T}^d$ we have that $X_t^1 - X_t^2$ satisfies the following stochastic differential equation for $s \in [0,t]$:

$$\begin{cases} d[X_s^1(x,\omega) - X_s^2(x,\omega)] = a(s,x,\omega)ds + \sigma dW_s(\omega) \\ X_0^1(x,\omega) - X_0^2(x,\omega) = 0 \end{cases}, \tag{6.2.1}$$

where

$$a(s,x,\omega) := b_1(s, X_s^1(x,\omega)) - b_2(s, X_s^2(x,\omega)) \quad \text{and} \quad \sigma := \sqrt{2\nu_1} - \sqrt{2\nu_2}.$$

We define the function $q_\delta(y) := \log\left(1 + \frac{|y|^2}{\delta^2}\right)$, and Q^δ as:

$$Q^\delta(t,x,\omega) := q_\delta(X_t^1(x,\omega) - X_t^2(x,\omega)) = \log\left(1 + \frac{|X_t^1(x,\omega) - X_t^2(x,\omega)|^2}{\delta^2}\right). \tag{6.2.2}$$

By Itô's lemma we have that Q^δ satisfies the following stochastic differential equation:

$$\begin{cases} dQ^\delta(t,x,\omega) &= \Big[a(t,x,\omega)\cdot\nabla_y q_\delta\big(X_t^1(x,\omega) - X_t^2(x,\omega)\big) + \frac{\sigma^2}{2}\Delta_y q_\delta\big(X_t^1(x,\omega) - X_t^2(x,\omega)\big)\Big]dt \\ & \quad + \Big[\sigma\Big(\sum_{i=1}^d \partial_{y_i} q_\delta\big(X_t^1(x,\omega) - X_t^2(x,\omega)\big)\Big)\Big]dW_t \\ Q^\delta(0,x,\omega) &= 0 \end{cases}. \tag{6.2.3}$$

Taking the expected value on both sides of (6.2.3) and integrating over $[0,t] \times \mathbb{T}^d$ we end up with

$$\begin{aligned} \int_{\mathbb{T}^d} \mathbb{E}_\omega[Q^\delta(t,x,\omega)]dx = &\int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega\left[a(s,x,\omega)\cdot\nabla_y q_\delta(X_s^1(x,\omega) - X_s^2(x,\omega))\right] dxds \\ &+ \frac{\sigma^2}{2}\int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega\left[\Delta q_\delta(X_s^1(x,\omega) - X_s^2(x,\omega))\right] dxds, \end{aligned}$$

where we use that the stochastic integral in (6.2.3) is a martingale with 0 as expected value. We can easily check that

$$\left|\nabla_y \log\left(1 + \frac{|y|^2}{\delta^2}\right)\right| \lesssim \frac{1}{\delta + |y|} \text{ and } \left|\Delta_y \log\left(1 + \frac{|y|^2}{\delta^2}\right)\right| \lesssim \frac{1}{\delta^2 + |y|^2}, \tag{6.2.4}$$

so

$$\int_{\mathbb{T}^d} \mathbb{E}_\omega[Q^\delta(t,x,\omega)]dx \lesssim \sigma^2 \frac{t}{2\delta^2} + \int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega \left[\frac{|a(s,x,\omega)|}{\delta + |X_s^1(x,\omega) - X_s^2(x,\omega)|} \right] dxds.$$

Since $a(t,x,\omega) = b_1(t, X_t^1(x,\omega)) - b_2(t, X_t^2(x,\omega))$, by triangular inequality:

$$|a(t,x,\omega)| \le |b_1(t, X_t^1(x,\omega)) - b_2(t, X_t^1(x,\omega))| + |b_2(t, X_t^1(x,\omega)) - b_2(t, X_t^2(x,\omega))|,$$

so

$$\begin{aligned} \int_{\mathbb{T}^d} \mathbb{E}_\omega[Q^\delta(t,x,\omega)]dx \lesssim \sigma^2 \frac{t}{\delta^2} &+ \int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega \left[\frac{|b_1(s, X_s^1(x,\omega)) - b_2(s, X_s^1(x,\omega))|}{\delta + |X_s^1(x,\omega) - X_s^2(x,\omega)|} \right] dxds \\ &+ \int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega \left[\frac{|b_2(s, X_s^1(x,\omega)) - b_2(s, X_s^2(x,\omega))|}{\delta + |X_s^1(x,\omega) - X_s^2(x,\omega)|} \right] dxds. \end{aligned} \tag{6.2.5}$$

For the first integral, use the measure-preserving property of the stochastic flow to get:

$$\begin{aligned} &\int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega \left[\frac{|b_1(s, X_s^1(x,\omega)) - b_2(s, X_s^1(x,\omega))|}{\delta + |X_s^1(x,\omega) - X_s^2(x,\omega)|} \right] dxds \\ &= \int_0^t \mathbb{E}_\omega \left[\int_{\mathbb{T}^d} \frac{|b_1(s, X_s^1(x,\omega)) - b_2(s, X_s^1(x,\omega))|}{\delta + |X_s^1(x,\omega) - X_s^2(x,\omega)|} dx \right] ds \\ &= \int_0^t \mathbb{E}_\omega \left[\int_{\mathbb{T}^d} \frac{|b_1(s,x) - b_2(s,x)|}{\delta + |X_s^1(x,\omega) - X_s^2(x,\omega)|} dx \right] ds \le \frac{\|b_1 - b_2\|_{L^1((s,t);L^1(\mathbb{T}^d))}}{\delta}. \end{aligned}$$

For the second integral, we use Lemma 4.1.2 to get:

$$\begin{aligned} &\int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega \left[\frac{|b_2(s, X_s^1(x,\omega)) - b_2(s, X_s^2(x,\omega))|}{\delta + |X_s^1(x,\omega) - X_s^2(x,\omega)|} \right] dxds \\ &\lesssim \int_0^t \int_{\mathbb{T}^d} \mathbb{E}_\omega \left[M(|Db_2|)(s, X_s^1(x,\omega)) + M(|Db_2|)(s, X_s^2(x,\omega)) \right] dxds \\ &= \int_0^t \mathbb{E}_\omega \left[\int_{\mathbb{T}^d} M(|Db_2|)(s, X_s^1(x,\omega)) + M(|Db_2|)(s, X_s^2(x,\omega)) dx \right] ds \\ &= 2 \int_0^t \int_{\mathbb{T}^d} M(|Db_2|)(s,x) dx d\tau \lesssim \|\nabla b_2\|_{L^1((0,t);L^p(\mathbb{T}^d))}, \end{aligned}$$

where we used the measure-preserving properties of the stochastic flow, Hölder's inequality and Lemma 4.1.1. In the end we have:

$$\int_{\mathbb{T}^d} \mathbb{E}_\omega[Q^\delta(t,x,\omega)]dx \lesssim \sigma^2 \frac{t}{\delta^2} + \frac{\|b_1 - b_2\|_{L^1(0,t;L^1(\mathbb{T}^d))}}{\delta} + \|\nabla b_2\|_{L^1(0,t;L^p(\mathbb{T}^d))} = K(t,\delta) = K.$$

By Chebyshev's inequality, there is a set $A_t \subset \mathbb{T}^d \times \Omega$ such that if $(x,\omega) \in A_t$:

$$Q^\delta(t,x,\omega) \le K/\eta \implies |X_t^1(x,\omega) - X_t^2(x,\omega)| \le \delta \exp\left(\frac{K}{\eta}\right)$$

and $(\mathscr{L}^d \otimes \mathrm{P})(A_t^c) \le \eta$. Then we have:

$$\int_{\mathbb{T}^d} \mathbb{E}_\omega[|X_t^1(x,\omega) - X_t^2(x,\omega)|]dx = \int_{A_t} |X_t^1(x,\omega) - X_t^2(x,\omega)|dxd\,\mathrm{P}(\omega)$$
$$+ \int_{A_t^c} |X_t^1(x,\omega) - X_t^2(x,\omega)|dxd\,\mathrm{P}(\omega) \lesssim \delta \exp\left(\frac{K}{\eta}\right) + \eta,$$

where we used that P is a probability measure and that the distance on $\mathbb{T}^d$ is bounded. Choosing $\eta = 2K|\log(\delta)|^{-1}$ we end up with

$$\int_{\mathbb{T}^d} \mathbb{E}_\omega[|X_t^1(x,\omega) - X_t^2(x,\omega)|]dx \lesssim \sqrt{\delta} + 2K|\log(\delta)|^{-1},$$

and finally choosing

$$\delta := \sqrt{t|\nu_1 - \nu_2|} + \|b_1 - b_2\|_{L^1(0,t;L^p(\mathbb{T}^d))},$$

we have that $K = K(t)$ is uniformly bounded and thus

$$\int_{\mathbb{T}^d} \mathbb{E}_\omega[|X_t^1(x,\omega) - X_t^2(x,\omega)|]dx$$
$$\lesssim \sqrt[4]{t|\nu_1 - \nu_2|} + \sqrt{\|b_1 - b_2\|_{L^1(0,t;L^p(\mathbb{T}^d))}} + \frac{\|\nabla b_2\|_{L^1(0,T;L^p(\mathbb{T}^d))}}{\left|\log\left(\sqrt{t|\nu_1 - \nu_2|} + \|b_1 - b_2\|_{L^1(0,t;L^p(\mathbb{T}^d))}\right)\right|}.$$

□

Proposition 6.2.2. ***Quantitative stability for solutions of*** **(ADE)**
Consider u_1, u_2 solutions of (ADE) *with respect to the same bounded initial datum u_0 but with different diffusivities ν_1, ν_2 and velocity fields b_1, b_2 which satisfy* (6.1.1). *Then :*

$$\sup_{t\in[0,T]} \mathrm{W}_1(u_1(t,\cdot), u_2(t,\cdot)) \lesssim C_t \|u_0\|_\infty \left(\sqrt[4]{|\nu_1 - \nu_2|} + \sqrt{\|b_1 - b_2\|_{L^1(s,t;L^p(\mathbb{T}^d))}} \right.$$
$$\left. + \frac{1}{\left|\log\left(\sqrt{|\nu_1 - \nu_2|} + \|b_1 - b_2\|_{L^1(0,t;L^p(\mathbb{T}^d))}\right)\right|} \right).$$

Proof. Identical to the proof of Theorem 4.3.1. We just need to use the representation (F-K), Kantorovich-Rubinstein duality (see Theorem 5.3.1) and Proposition 6.2.1. □

We now give a series of results on stochastic flows which are practically identical to their deterministic counterparts we saw in Chapter 4 and in Chapter 5. The reason why we can essentially copy the proofs of the deterministic case is that under our assumptions such results are "fundamentally deterministic". Indeed, since we consider an additive noise, each time we subtract two stochastic flows at fixed $\omega \in \Omega$ (i.e. for a fixed realisation of the Brownian motion W_t) the noise cancels out. We also rely crucially on the fact that, under our assumptions, the stochastic flows are almost surely (with respect to ω) measure preserving [78, Lemma 9].

Theorem 6.2.1.
Let b satisfy (6.1.1) *and let X_t^ν be the stochastic flow associated to b with diffusivity $\nu > 0$. We define the following integral quantity:*

$$A_p(R, X^\nu, \omega) := \left[\int_{B_R(0)} \left(\sup_{0\le t\le T} \sup_{0<r<2R} \fint_{B_r(x)} \log\left(1 + \frac{|X_t^\nu(x,\omega) - X_t^\nu(y,\omega)|^2}{r^2}\right) dy \right)^p dx \right]^{1/p}. \tag{6.2.6}$$

We have:

$$A_p(R, X^\nu, \omega) \le C(p, d, R, L, \|D_x b\|_{L^1(0,T;L^p(\mathbb{T}^d))}) \quad \text{for almost every } \omega \in \Omega,$$

where the value of C does not depend on ω. Since we assumed to work in a torus, if R is big enough we have that $\mathbb{T}^d = B_R(x)$ for every $x \in \mathbb{T}^d$, so we can also define

$$A_p(X^\nu, \omega) := \inf_{\{R \mid \mathbb{T}^d \subset B_R(x)\ \forall x \in \mathbb{T}^d\}} A_p(R, X^\nu, \omega),$$

for which the same estimate holds.

Proof. For $0 \le t \le T$, $0 < r < 2R$ and $x \in B_R(0)$ we define the quantity

$$Q(t, x, y, r, \omega) := \log\left(1 + \frac{|X_t^\nu(x, \omega) - X_t^\nu(y, \omega)|^2}{r^2}\right).$$

We argue as in Proposition 6.2.1. Since the noise cancels by considering the difference of the stochastic flows, we have that by (6.2.3) and the bounds in (6.2.4):

$$Q(t, x, y, r, \omega) \le Q(0, x, y, r, \omega) + \int_0^t \frac{|b(s, X_s^\nu(x, \omega)) - b(s, X_s^\nu(y, \omega))|}{r + |X_s^\nu(x, \omega) - X_s^\nu(y, \omega)|} ds.$$

Considering the average with respect to the y variable over a ball $B_r(x)$ we have:

$$\fint_{B_r(x)} Q(t, x, y, r, \omega) dy \lesssim c + \int_0^t \fint_{B_r(x)} \frac{|b(s, X_s^\nu(x, \omega)) - b(s, X_s^\nu(y, \omega))|}{r + |X_s^\nu(x, \omega) - X_s^\nu(y, \omega)|} dy ds,$$

and by Lemma 4.1.2

$$\fint_{B_r(x)} Q(t, x, y, r, \omega) dy \lesssim c + \int_0^t MDb(s, X_s^\nu(x, \omega)) ds + \int_0^t \fint_{B_r(x)} MDb(s, X_s^\nu(y, \omega)) dy ds.$$

Taking the supremum in r and t:

$$\begin{aligned} &\sup_{0<t<T} \sup_{0<r<2R} \fint_{B_r(x)} Q(t, x, y, r, \omega) dy \\ &\lesssim c + \int_0^T MDb(s, X_s^\nu(x, \omega)) ds + \int_0^T \sup_{0<r<2R} \fint_{B_r(x)} MDb(s, X_s^\nu(y, \omega)) dy ds. \end{aligned}$$

We conclude taking the $L^p(\mathbb{T}^d)$ norm with respect to the x variable on both sides and arguing as in Theorem 4.2.1, using the fact that for a.e. $\omega \in \Omega$ the stochastic flow is measure preserving to change variables when needed, and using Hölder's inequality and Lemma 4.1.1. □

Proposition 6.2.3. ***Almost everywhere Lipschitz continuity of the stochastic flow outside a small set***
Let $X^\nu : [0, T) \times \mathbb{T}^d \to \mathbb{T}^d$ be a stochastic flow, as in Definition 6.1.1. Then, for almost every $\omega \in \Omega$, for every $\varepsilon > 0$ and every $R > 0$ we can find a set (depending on $\omega \in \Omega$) $K_\omega \subset B_R(0)$ such that $|B_R(0) \setminus K_\omega| \le \varepsilon$ and for any $0 \le t \le T$ it holds

$$\mathrm{Lip}(X_t^\nu(\cdot, \omega)|_{K_\omega}) \le \exp\left(\frac{A_p(X^\nu, \omega)}{\varepsilon^{1/p}}\right).$$

Proof. Identical to Proposition 4.2.1, up to considering a square root in the final step. □

6.3 Euler-Maruyama scheme and applications to the advection-diffusion equation

We now define our approximation scheme for (SDE), which is known in the literature as the *Euler-Maruyama* method. We will often omit explicitly writing the dependence of the stochastic flow on ω, writing $X^\nu(t,x)$ in place of $X_t^\nu(x,\omega)$.

Definition 6.3.1. ***Euler-Maruyama method***
Let $b \in [L^1(0,T;W^{1,p}(\mathbb{T}^d)) \cap L^\infty((0,T)\times\mathbb{T}^d)]^d$ with $1 < p \le d$, and let

$$b_\delta(t,\cdot) := \eta_\delta * b(t,\cdot) \tag{6.3.1}$$

for some $\delta > 0$, where η is a standard mollifier in $\mathbb{R}^d$, $\eta_\delta(x) := \delta^{-d}\eta(x/\delta)$ and the convolution is in the spatial variables only. We denote by X_{EM} (where "EM" stands for "Euler-Maruyama") the Euler-Maruyama approximation of the stochastic flow X_t^ν of b with diffusivity $\nu > 0$, with time step Δt and regularisation parameter δ. For $t \in [t_n, t_{n+1}]$ we define:

$$X_{EM}(t,x,\omega) := X_{EM}(t_n,x,\omega)+(t-t_n)\fint_{t_n}^{t_{n+1}} b_\delta(s,X_{EM}(t_n,x,\omega))ds+(t-t_n)\sqrt{2\nu}(W_{t_{n+1}}(\omega)-W_{t_n}(\omega)), \tag{6.3.2}$$

where $0 = t_0 < t_1 < \cdots < t_M = T$ is a partition of $[0,T]$ with $t_{i+1}-t_i = \Delta t$ and $X_{EM}(0,x,\omega) = x\ \forall x \in \mathbb{T}^d$ and $\forall \omega \in \Omega$. If $p > d$, we define X_{EM} directly with the vector field b, without any prior regularisation, as

$$X_{EM}(t,x,\omega) := X_{EM}(t_n,x,\omega)+(t-t_n)\fint_{t_n}^{t_{n+1}} b(s,X_{EM}(t_n,x,\omega))ds+(t-t_n)\sqrt{2\nu}(W_{t_{n+1}}(\omega)-W_{t_n}(\omega)). \tag{6.3.3}$$

We remark that, in order to keep the notation as light as possible, we omit explicitly writing the dependence of Φ_E on $\Delta t, \omega$ and (eventually) δ.

Proposition 6.3.1. ***Compressibility estimate and injectivity of X_{EM}***
Consider X_{EM} defined as in Definition 6.3.1 with $1 < p \le d$ and with $\delta = \sqrt{\Delta t}$. Then for a.e. $\omega \in \Omega$ there exists a constant $c > 0$ not dependent on Δt and on ω such that

$$\det(\nabla_x X_{EM}(t,x)) > c \quad \text{for every } (t,x) \in [0,T]\times\mathbb{T}^d.$$

Moreover, $X_{EM}(t,\cdot)$ is Lipschitz continuous for every t, and in particular it is injective.

Proof. Since the noise is additive, the proofs of the compressibility bound and of the Lipschitz continuity (and hence of injectivity) follow exactly as in Proposition 5.2.1 and Proposition 5.2.2. □

Proposition 6.3.2. ***L^p stability estimate for X_δ and X_{EM}***
Let X_{EM} be defined as in (6.3.1) *with b satisfying assumptions* (6.1.1) *with $1 < p \le d$, and let X^ν, X_δ^ν be the stochastic flows of b and b_δ respectively, with $\delta = \sqrt{\Delta t}$. Then:*

$$\mathbb{E}_\omega[\|X_\delta^\nu(t,\cdot) - X_{EM}(t,\cdot)\|_{L^p(\mathbb{T}^d)}] \lesssim C_{t,\nu\Delta t}|\log(\Delta t)|^{-1} \text{ for every } t \in [0,T),$$

where X_δ is the flow of b_δ. If $b \in L^1(0,T;W^{1,p}(\mathbb{T}^d))$ with $p > d$, then we consider X_{EM} defined as in (6.3.3) *and it holds*

$$\mathbb{E}_\omega[\|X^\nu(t,\cdot) - X_{EM}(t,\cdot)\|_{L^p(\mathbb{T}^d)}] \lesssim C_{t,\nu\Delta t}|\log(\Delta t)|^{-1} \text{ for every } t \in [0,T).$$

In both cases, $C_{t,\nu\Delta t}$ is a positive constant such that $C_{t,\nu\Delta t} \to 0$ as $t \to 0$ for $\nu\Delta t$ fixed, and $C_{t,\nu\Delta t} \approx (1+\sqrt{\nu\Delta t})$ at t fixed.

Proof. In the first part of the proof we will work without loss of generality with X_δ^ν with the assumption $1 < p \le d$, since the argument is identical if $p > d$. The difference will appear only at the end of the proof, where we differentiate the two cases. We want to prove that

$$\mathbb{E}_\omega\left[\left\|\log\left(1+\frac{|X_\delta^\nu(t,\cdot)-X_{EM}(t,\cdot)|}{\Delta t}\right)\right\|_{L^p(\mathbb{T}^d)}\right] \le C_t,$$

with a constant C_t not dependent on Δt such that $C_t \to 0$ as $t \to 0$. Once we prove that, the conclusion follows as in the last part of the proof of Proposition 6.2.1, i.e. using Chebyshev's inequality. Notice that, by the concavity of the logarithm and by the fact that $\log(1+x)$ is an increasing monotone function, for every $a, b \ge 0$ it holds

$$\log(1+b)-\log(1+a) \le \frac{d}{dx}\Big[\log(1+x)\Big]_{x=a}|b-a| = \frac{|b-a|}{1+a}.$$

In particular, let $t = t_n$ for simplicity. We have (since the estimate is pointwise we omit writing the dependence on x to keep the notation as light as possible):

$$\begin{aligned}
&\log\left(1+\frac{|X_\delta^\nu(t_n)-X_{EM}(t_n)|}{\Delta t}\right)-\log\left(1+\frac{|X_\delta^\nu(t_{n-1})-X_{EM}(t_{n-1})|}{\Delta t}\right)\\
&\le \frac{\Delta t}{|X_\delta^\nu(t_{n-1})-X_{EM}(t_{n-1})|+\Delta t}\,\frac{|\,|X_\delta^\nu(t_n)-X_{EM}(t_n)|-|X_\delta^\nu(t_{n-1})-X_{EM}(t_{n-1})|\,|}{\Delta t}\\
&\le \frac{|X_\delta^\nu(t_n)-X_\delta^\nu(t_{n-1})-(X_{EM}(t_n)-X_{EM}(t_{n-1}))|}{|X_\delta^\nu(t_{n-1})-X_{EM}(t_{n-1})|+\Delta t}\\
&= \frac{|\int_{t_{n-1}}^{t_n} b_\delta(s,X_\delta^\nu(s))-b_\delta(s,X_{EM}(t_{n-1}))ds|}{|X_\delta^\nu(t_{n-1})-X_{EM}(t_{n-1})|+\Delta t}.
\end{aligned}$$

Iterating the estimate above we get, by telescopic summing:

$$\begin{aligned}
&\log\left(1+\frac{|X_\delta^\nu(t_n)-X_{EM}(t_n))|}{\Delta t}\right) \le \sum_{j=0}^{n-1}\frac{|\int_{t_j}^{t_{j+1}} b_\delta(s,X_\delta^\nu(s))-b_\delta(s,X_{EM}(t_j))ds|}{|X_\delta^\nu(t_j)-X_{EM}(t_j)|+\Delta t} \qquad (6.3.4)\\
&\le \sum_{j=0}^{n-1}\frac{|\int_{t_j}^{t_{j+1}} b_\delta(s,X_\delta^\nu(s))-b_\delta(s,X_\delta^\nu(t_j))ds|+|\int_{t_j}^{t_{j+1}} b_\delta(s,X_\delta^\nu(t_j))-b_\delta(s,X_{EM}(t_j))ds|}{|X_\delta^\nu(t_j)-X_{EM}(t_j)|+\Delta t}\\
&\le \underbrace{\sum_{j=0}^{n-1}\frac{\int_{t_j}^{t_{j+1}} |b_\delta(s,X_\delta^\nu(s))-b_\delta(s,X_\delta^\nu(t_j))|ds}{|X_\delta^\nu(t_j)-X_{EM}(t_j)|+\Delta t}}_{=1}+\underbrace{\sum_{j=0}^{n-1}\frac{\int_{t_j}^{t_{j+1}} |b_\delta(s,X_\delta^\nu(t_j))-b_\delta(s,X_{EM}(t_j))|ds}{|X_\delta^\nu(t_j)-X_{EM}(t_j)|+\Delta t}}_{=2}.
\end{aligned}$$

Now, notice that by Lemma 4.1.2:

$$\begin{aligned}
&\int_{t_j}^{t_{j+1}} |b_\delta(s,X_\delta^\nu(s))-b_\delta(s,X_\delta^\nu(t_j))|ds\\
&\lesssim \int_{t_j}^{t_{j+1}} |X_\delta^\nu(s)-X_\delta^\nu(t_j)|\;[M(|Db_\delta|)(s,X_\delta^\nu(s))+M(|Db_\delta|)(s,X_\delta^\nu(t_j))]ds\\
&= \int_{t_j}^{t_{j+1}} \left|\int_{t_j}^{s} b_\delta(s,X_\delta^\nu(s))+\sqrt{2\nu}(W_s-W_{t_j})ds\right|[M(|Db_\delta|)(s,X_\delta^\nu(s))+M(|Db_\delta|)(s,X_\delta^\nu(t_j))]ds\\
&\le \left(\|b_\delta\|_\infty\Delta t+\sqrt{2\nu}\int_{t_j}^{t_{j+1}}|W_\tau-W_{t_j}|d\tau\right)\int_{t_j}^{t_{j+1}}[M(|Db_\delta|)(s,X_\delta^\nu(s))+M(|Db_\delta|)(s,X_\delta^\nu(t_j))]ds.
\end{aligned}$$

We now consider the L^p norm and we change variables to get:

$$① \lesssim \sum_{j=0}^{n-1} \left(\|b_\delta\|_\infty + \frac{\sqrt{2\nu} \int_{t_j}^{t_{j+1}} |W_\tau - W_{t_j}| d\tau}{\Delta t} \right) \int_{t_j}^{t_{j+1}} \|M(|Db_\delta|)(s, \cdot)\|_{L^p(\mathbb{T}^d)} ds. \tag{6.3.5}$$

Considering the expected value (with respect to ω), the only term which still depends on ω is $|W_\tau - W_{t_j}|$. It is a well-known fact that, up to a dimensional constant, $\mathbb{E}_\omega[|W_\tau - W_{t_j}|] \approx \sqrt{\tau - t_j}$. So:

$$\int_{t_j}^{t_{j+1}} \mathbb{E}_\omega[|W_\tau - W_{t_j}|] d\tau \approx \int_{t_j}^{t_{j+1}} \sqrt{\tau - t_j} d\tau \approx (\Delta t)^{3/2}.$$

So, we have

$$\mathbb{E}_\omega \left[\left\| ① \right\|_{L^p(\mathbb{T}^d)} \right] \lesssim (\|b_\delta\|_\infty + \sqrt{\nu \Delta t}) \|M(|Db_\delta|)\|_{L^1(0,t;L^p(\mathbb{T}^d))}.$$

We just need to apply Lemma 4.1.1 to conclude. As for ②, we need to differentiate the cases $1 < p \leq d$ and $p > d$.

- If $1 < p \leq d$, we repeat the argument above. We have

$$\begin{aligned} &\int_{t_j}^{t_{j+1}} |b_\delta(s, X_\delta^\nu(t_j)) - b_\delta(s, X_{EM}(t_j))| ds \\ &\lesssim |X_h^\nu(t_j) - X_{EM}(t_j)| \int_{t_j}^{t_{j+1}} [M(|Db_\delta|)(s, X_\delta^\nu(t_j)) + M(|Db_\delta|)(s, X_{EM}(t_j))] ds, \end{aligned}$$

  and we can estimate $|X_\delta^\nu(t_j) - X_{EM}(t_j)|/(|X_h^\nu(t_j) - X_{EM}(t_j)| + \Delta t) \leq 1$. Then, after summing over j:

$$② \lesssim \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} [M(|Db_\delta|)(s, X_\delta^\nu(t_j)) + M(|Db_\delta|)(s, X_{EM}(t_j))] ds.$$

  After considering the L^p norm and changing variables the estimate for the first addendum follows immediately. For the second addendum, we conclude in the same way using Proposition 6.3.1.

- If $p > d$, we recall that we work directly with X^ν, and not with X_δ^ν. In this case we use Lemma 4.1.3 with $\tilde{p} \in (d, p)$ to get:

$$\begin{aligned} \frac{\int_{t_j}^{t_{j+1}} |b(s, X^\nu(t_j)) - b(s, X_{EM}(t_j))| ds}{|X^\nu(t_j) - X_{EM}(t_j)| + \Delta t} &\lesssim \frac{|X^\nu(t_j) - X_{EM}(t_j)| \int_{t_j}^{t_{j+1}} f(s, X^\nu(t_j)) ds|}{|X^\nu(t_j) - X_{EM}(t_j)| + \Delta t} \\ &\leq \int_{t_j}^{t_{j+1}} f(s, X^\nu(t_j)) ds, \end{aligned} \tag{6.3.6}$$

  where $f(s,x) := [M(|Du|^{\tilde{p}})(s,x)]^{1/\tilde{p}} \in L^1(0,T; L^{\tilde{p}}(\mathbb{T}^d))$. To conclude we just need to sum over j, consider the L^p norm and use the compressibility bound on X^ν to change variables. Notice that $p/\tilde{p} > 1$, so we can apply Lemma 4.1.1.

At this point, we can conclude the proof using Chebyshev's inequality, as at the end of the proof of Proposition 6.2.1. □

We are now ready to give the main result of this section, i.e. a probabilistic approximation of the solution of (ADE) via the Euler-Maruyama method. The proof follows almost verbatim the one of Theorem 5.3.2.

Theorem 6.3.1. ***Singular probabilistic approximation of the solution of*** **(ADE)** ***in*** **1*-Wasserstein distance*** *Assume that b satisfies* (6.1.1) *and $u_0 \geq 0$ and bounded, and let u be the solution of* (ADE). *Assume that the domain $\mathbb{T}^d$ is partitioned into sets $\{Q_i\}_{i=1}^N$ such that $\operatorname{diam}(Q_i) \leq \Delta x$ $\forall i = 1, \dots, N$ for some space discretization scale $\Delta x > 0$. We choose, for every set Q_i, a random starting point $x_0^i \in Q_i$ according to a uniform probability distribution on Q_i and we define the weights $M_i := \int_{Q_i} u_0 dx$. For $t \in [0, T]$, consider the measure given by:*

$$u_{EM}(t,\cdot) := \sum_{i=1}^N M_i \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0^i,\omega^i)}] = \sum_{i=1}^N \mathbb{E}_\omega[X_{EM}(t,\cdot,\omega)_\#(M_i \delta_{x_0^i})], \tag{6.3.7}$$

where $\bar{x}_0 = (x_0^1, x_0^2, \dots, x_0^N)$ and X_{EM} is defined as in (6.3.2) *with a time step $\Delta t > 0$ and with $\delta = \sqrt{\Delta t}$ if $1 < p \leq d$ or as in* (6.3.3) *otherwise. Denoting with W_1 the 1-Wasserstein distance defined in* (5.3.2) *it holds that, in expected value with respect to the choice of the points $\bar{x}_0$, for every $\alpha \in (0,1)$:*

$$\mathbb{E}_{x_0}[\mathrm{W}_1(u(t,\cdot), u_{EM}(t,\bar{x}_0))] \lesssim (\Delta x)^{1-\alpha} + \frac{|\log(\Delta x)|^{-p}}{\alpha^p} + C_{t,\nu\Delta t}|\log(\Delta t)|^{-1}, \tag{6.3.8}$$

with $C_{t,\nu\Delta t} > 0$ as in Proposition 6.3.2. We can also estimate the variance of $\mathrm{W}_1(u(t,\cdot), u_{EM}(t,\bar{x}_0))$ if $N \lesssim (\Delta x)^{-d}$. In this case:

$$\operatorname{Var}_{x_0}[\mathrm{W}_1(u(t,\cdot), u_{EM}(t,\bar{x}_0))] \lesssim (\Delta x)^{2-2\alpha} + \frac{|\log(\Delta x)|^{-p}}{\alpha^p} + C_{t,\nu\Delta t}|\log(\Delta t)|^{-1}. \tag{6.3.9}$$

Proof. First of all, we notice that in order to prove (6.3.8) we can work without loss of generality with X_δ^ν instead of X^ν if $p > d$. By Proposition 6.2.1 and (F-K), using Kantorovich-Rubinstein duality, we have:

$$\begin{aligned} \mathrm{W}_1(\mathbb{E}_\omega[X^\nu(t,\cdot)_\# u_0], \mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\# u_0]) &\leq \int_{\mathbb{T}^d} u_0(x)\mathbb{E}_\omega[|X^\nu(t,x) - X_\delta^\nu(t,x)|]dx \qquad (6.3.10)\\ \leq \|u_0\|_{L^\infty}\mathbb{E}_\omega[\|X^\nu(t,\cdot) - X_\delta^\nu(t,\cdot)\|]_{L^1} &\lesssim C_{t,\nu\Delta t}|\log(\delta)|^{-1} \approx C_{t,\nu\Delta t}|\log(\Delta t)|^{-1}. \end{aligned}$$

At time t fixed, consider $\bar{x}_0 = (x_0^1, \dots, x_0^N)$ as a random variable in $((\mathbb{T}^d)^N, \mathcal{B}(\mathbb{T}^d)^N, \bar{P}_{x_0})$, with $\mathcal{B}(\mathbb{T}^d)$ the Borel sigma algebra on $\mathbb{T}^d$ and

$$\bar{\mathrm{P}}_{x_0} = \mathrm{P}_1 \otimes \mathrm{P}_2 \cdots \otimes \mathrm{P}_N, \quad \text{with } \mathrm{P}_i \text{ a uniform distribution over } Q_i. \tag{6.3.11}$$

We also define $u_0^i := u_0 \llcorner Q_i$ and $\bar{u}_0^i := (u_0 \llcorner Q_i)/M_i$ if $M_i > 0$, otherwise $M_i = \bar{u}_0^i = 0$, where $\mu \llcorner A$ denotes the restriction of a measure μ to the set A. In the following, each time we consider an expected value $\mathbb{E}_{x_0}[\cdot]$ or a variance $\operatorname{Var}_{x_0}[\cdot]$ it will be with respect to the probability $\bar{\mathrm{P}}_{x_0}$. We will denote with $\mathbb{E}_\omega[\cdot]$ the expected value with respect to the probability space $(\Omega, \mathcal{F}, P)$ the Brownian motion W_t is defined with respect to. We consider the measures

$$\mu_i := \mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#(u_0 \llcorner Q_i)] \text{ and } \nu_i := M_i\mathbb{E}_\omega[\delta_{X_{EM}(t,x_0^i)}] = \mathbb{E}_\omega[X_{EM}(t,\cdot)_\#(M_i\delta_{x_0^i})],$$

which have the same mass, so we can apply Lemma 5.3.1. For the moment we will work in a fixed set $Q = Q_i$ of the partition such that $M_i \neq 0$, so in order to keep the notation lighter we will drop the subscript and in the following we will implicitly assume that the measures are restricted to Q and normalized, also we will write x_0 instead of x_0^i. Since the solution u of (ADE) can be represented as $u(t,\cdot) = \mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0]$, we need to estimate $\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}])]$. By triangle inequality

$$\begin{aligned} \mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}])] \leq &\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0], \mathbb{E}_\omega[\delta_{X_\delta^\nu(t,x_0)}])] \qquad (6.3.12)\\ &+ \mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[\delta_{X_\delta^\nu(t,x_0)}], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}])]. \end{aligned}$$

For the first addendum in (6.3.12), recalling that $\int_Q \bar{u}_0(x)dx = 1$:

$$\begin{aligned}
&\mathrm{W}_1(\mathbb{E}_\omega[X^\nu_\delta(t,\cdot)_\# \bar{u}_0], \mathbb{E}_\omega[\delta_{X^\nu_\delta(t,x_0)}]) \qquad (6.3.13)\\
&= \sup_{\mathrm{Lip}(h)\le 1} \left(\int_Q \bar{u}_0(x)\mathbb{E}_\omega[h(X^\nu_\delta(t,x))]dx - \int_Q \bar{u}_0(x)\mathbb{E}_\omega[h(X^\nu_\delta(t,x_0))]dx \right)\\
&\le \|\bar{u}_0\|_\infty \int_Q \mathbb{E}_\omega[|X^\nu_\delta(t,x) - X^\nu_\delta(t,x_0)|].
\end{aligned}$$

Using Proposition 6.2.3, for almost every $\omega \in \Omega$ we can consider a set $K_\omega \subset \mathbb{T}^d$ such that $|K^c_\omega| = \alpha^{-p} A_p(R, X^\nu_\delta)^p |\log(\Delta x)|^{-p}$ for $\alpha \in (0,1)$ such that $\mathrm{Lip}(X^\nu_\delta(t,\cdot) \llcorner K_\omega) \le (\Delta x)^{-\alpha}$ for every $0 \le t \le T$. Notice that $A_p(R, X_\delta)$ does not depend on the mollification scale δ, since $\|D_x b_\delta\|_{L^1(L^p)} \le \|D_x b\|_{L^1(L^p)}$ for every $\delta > 0$ by Young's convolution inequality. Then using that $|x - x_0| \le \mathrm{diam}(Q) \le \Delta x$ and (6.3.13):

$$\begin{aligned}
&\mathbb{E}_{x_0}\left[\mathrm{W}_1(\mathbb{E}_\omega[X^\nu_\delta(t,\cdot)_\# \bar{u}_0], \mathbb{E}_\omega[\delta_{X^\nu_\delta(t,x_0)}])\right] \qquad (6.3.14)\\
&\le \frac{1}{\mathscr{L}^d(Q)}\mathbb{E}_\omega\left[\int_{Q\cap K_\omega}\left(\int_{Q\cap K_\omega} \bar{u}_0(x)|X^\nu_\delta(t,x) - X^\nu_\delta(t,x_0)|dx \right) dx_0 \right.\\
&+ \int_{Q\cap K_\omega}\left(\int_{Q\cap K^c_\omega} \bar{u}_0(x)|X^\nu_\delta(t,x) - X^\nu_\delta(t,x_0)|dx \right) dx_0\\
&\left. + \int_{Q\cap K^c_\omega}\left(\int_{Q} \bar{u}_0(x)|X^\nu_\delta(t,x) - X^\nu_\delta(t,x_0)|dx \right) dx_0 \right]\\
&\lesssim \mathbb{E}_\omega\left[(\Delta x)^{1-\alpha}\|\bar{u}_0\|_{L^1(Q\cap K_\omega)} + \left(\mathscr{L}^d(Q\cap K^c_\omega) + \frac{\mathscr{L}^d(Q\cap K^c_\omega)}{\mathscr{L}^d(Q)} \right) \right]\\
&\lesssim \mathbb{E}_\omega\left[(\Delta x)^{1-\alpha}\|\bar{u}_0\|_{L^1(Q\cap K_\omega)} + \frac{\mathscr{L}^d(Q\cap K^c_\omega)}{\mathscr{L}^d(Q)} \right],
\end{aligned}$$

where we used that for some $C > 0$ (since we are in a torus and $\bar{u}_0 \in L^\infty$):

$$\bar{u}_0\mathbb{E}_\omega[|X^\nu_\delta(t,x) - X_{EM}(t,x_0)|] \le C \text{ and that } \|\bar{u}_0\|_{L^1(Q)} = 1 \text{ by construction.}$$

As for the other addendum in (6.3.12), we have:

$$\mathrm{W}_1(\mathbb{E}_\omega[\delta_{X^\nu_\delta(t,x_0)}], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}]) \le \mathbb{E}_\omega[|X^\nu_\delta(t,x_0) - X_{EM}(t,x_0)|]. \qquad (6.3.15)$$

So, considering the expected value with respect to x_0:

$$\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[\delta_{X^\nu_\delta(t,x_0)}], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}])] \le \frac{1}{\mathscr{L}^d(Q)}\int_Q \mathbb{E}_\omega[|X^\nu_\delta(t,x_0) - X_{EM}(t,x_0)|]dx_0. \qquad (6.3.16)$$

Then using Lemma 5.3.1:

$$\begin{aligned}
&\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[X^\nu_\delta(t,\cdot)_\# u_0], u_{EM}(t,\bar{x}_0))] \le \mathbb{E}_{x_0}\left[\sum_{i=1}^N M_i \mathrm{W}_1\left(\mathbb{E}_\omega[X^\nu_\delta(t,\cdot)_\# \bar{u}^i_0], \mathbb{E}_\omega[\delta_{X_{EM}(t,x^i_0)}] \right) \right] \qquad (6.3.17)\\
&= \sum_{i=1}^N M_i \mathbb{E}_{x_0}\left[\mathrm{W}_1\left(\mathbb{E}_\omega[X^\nu_\delta(t,\cdot)_\# \bar{u}^i_0], \mathbb{E}_\omega[\delta_{X_{EM}(t,x^i_0)}] \right) \right]\\
&\lesssim \sum_{i=1}^N M_i \mathbb{E}_{x_0}\mathbb{E}_\omega\left[(\Delta x)^{1-\alpha}\|\bar{u}^i_0\|_{L^1(Q_i\cap K_\omega)} + \frac{\mathscr{L}^d(Q\cap K^c_\omega)}{\mathscr{L}^d(Q)} + \frac{1}{\mathscr{L}^d(Q_i)}\int_{Q_i} |X_{EM}(t,x_0) - X^\nu_\delta(t,x_0)|dx_0 \right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} \mathbb{E}_{x_0}\mathbb{E}_{\omega}\left[(\Delta x)^{1-\alpha}\|u_0^i\|_{L^1(Q_i\cap K_\omega)} + \|u_0\|_\infty \mathscr{L}^d(Q_i\cap K_\omega^c) + \|u_0\|_\infty \int_{Q_i} |X_{EM}(t,x_0) - X_\delta^\nu(t,x_0)|dx_0\right] \\
&\lesssim (\Delta x)^{1-\alpha}\|u_0\|_{L^1(\mathbb{T}^d)} + \|u_0\|_\infty \mathbb{E}_\omega[\mathscr{L}^d(K_\omega^c)] + \|u_0\|_\infty \mathbb{E}_\omega[\|X_\delta^\nu(t,x_0) - X_{EM}(t,x_0)\|_{L^1(\mathbb{T}^d)}] \\
&\lesssim (\Delta x)^{1-\alpha} + \|u_0\|_\infty\left(\frac{|\log(\Delta x)|^{-p}}{\alpha^p} + C_{t,\nu\Delta t}|\log(\Delta t)|^{-1}\right),
\end{aligned}
$$

where we have used that $\|u_0\|_{L^1(K_\omega^c)} \le \|u_0\|_\infty \mathscr{L}^d(K_\omega^c)$ with $\mathscr{L}^d(K_\omega^c) \approx \alpha^{-p}|\log(\Delta x)|^{-p}$, independent of ω, and that by Proposition 6.3.2

$$
\|X_\delta^\nu(t,\cdot) - X_{EM}(t,\cdot)\|_{L^1(\mathbb{T}^d)} \lesssim C_{t,\nu\Delta t}|\log(\Delta t)|^{-1}. \tag{6.3.18}
$$

We now estimate the variance. By definition

$$
\mathrm{Var}_{x_0}[X] = \mathbb{E}_{x_0}[X^2] - \mathbb{E}_{x_0}[X]^2.
$$

We will simply estimate $\mathrm{Var}_{x_0}(X)$ from above with the second moment $\mathbb{E}_{x_0}[X^2]$. First of all we verify that we can work without loss of generality with X_δ^ν instead of X^ν. Indeed by (6.3.10) we have:

$$
\mathrm{W}_1(\mathbb{E}_\omega[X^\nu(t,\cdot)_\#\bar{u}_0], \mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0])^2 \lesssim C_{t,\nu\Delta t}^2|\log(\Delta t)|^{-2}. \tag{6.3.19}
$$

We now restrict to working in a fixed set Q_i. As we did before, we omit writing the index i and we will denote $(u_0 \llcorner Q_i)/M_i$ and x_0^i as $\bar{u}_0$ and x_0. It holds:

$$
\begin{aligned}
\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}])^2 \le\ & 2\,\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0], \mathbb{E}_\omega[\delta_{X_\delta^\nu(t,x_0)}])^2 \\
&+ 2\,\mathrm{W}_1(\mathbb{E}_\omega[\delta_{X_\delta^\nu(t,x_0)}], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}])^2.
\end{aligned} \tag{6.3.20}
$$

For the first addendum in the right-hand side of (6.3.20), using (6.3.13), the same choice of K_ω, the decomposition in (6.3.14) and the elementary inequality $(a+b)^2 \le 2a^2 + 2b^2$:

$$
\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0], \mathbb{E}_\omega[\delta_{X_\delta^\nu(t,x_0)}])^2] \lesssim \mathbb{E}_\omega\left[(\Delta x)^{2-2\alpha}\|\bar{u}_0\|_{L^1(Q\cap K_\omega)}^2 + \frac{\mathscr{L}^d(Q\cap K_\omega^c)^2}{\mathscr{L}^d(Q)^2}\right]. \tag{6.3.21}
$$

For the second addendum in (6.3.20), considering the expected value of (6.3.15) with respect to x_0:

$$
\begin{aligned}
&\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[\delta_{X_\delta^\nu(t,x_0)}], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0)}])^2] \le \frac{1}{\mathscr{L}^d(Q)}\int_Q \mathbb{E}_\omega[|X_\delta^\nu(t,x_0) - X_{EM}(t,x_0)|^2]dx_0 \qquad (6.3.22)\\
&\le \frac{\|X_\delta^\nu(t,\cdot) - X_{EM}(t,\cdot)\|_\infty}{\mathscr{L}^d(Q)}\int_Q \mathbb{E}_\omega[|X_\delta^\nu(t,x_0) - X_{EM}(t,x_0)|]dx_0 \\
&\lesssim \frac{1}{\mathscr{L}^d(Q)}\int_Q \mathbb{E}_\omega[|X_\delta^\nu(t,x_0) - X_{EM}(t,x_0)|]dx_0.
\end{aligned}
$$

The conclusion follows as in (6.3.17), using (6.3.21), (6.3.22) and that by Cauchy-Schwartz:

$$
\mathrm{W}_1(\mu,\nu) \le \sum_{i=1}^N M_i\,\mathrm{W}_1(\bar{\mu}_i,\bar{\nu}_i) \implies \mathrm{W}_1(\mu,\nu)^2 \le N\sum_{i=1}^N M_i^2\,\mathrm{W}_1(\bar{\mu}_i,\bar{\nu}_i)^2.
$$

Indeed

$$
\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\# u_0], u_{EM}(t,\bar{x}_0))^2] \le N\sum_{i=1}^N M_i^2\mathbb{E}_{x_0}\left[\mathrm{W}_1\left(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\#\bar{u}_0^i], \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0^i)}]\right)^2\right] \tag{6.3.23}
$$

$$\lesssim N\sum_{i=1}^{N} M_i^2 \mathbb{E}_\omega\left[(\Delta x)^{2-2\alpha}\|\bar{u}_0^i\|^2_{L^1(Q_i\cap K_\omega)} + \frac{\mathscr{L}^d(Q_i\cap K_\omega^c)^2}{\mathscr{L}^d(Q_i)^2} + \frac{1}{\mathscr{L}^d(Q_i)}\int_{Q_i}|X_{EM}(t,x_0)-X_\delta^\nu(t,x_0)|dx_0\right]$$
$$\lesssim N\sum_{i=1}^{N} M_i \mathbb{E}_\omega\left[(\Delta x)^{2-2\alpha}\|u_0^i\|_{L^1(Q_i)} + \mathscr{L}(K_\omega^c\cap Q_i) + \|u_0\|_\infty\int_{Q_i}|X_{EM}(t,x_0)-X_\delta^\nu(t,x_0)|dx_0\right].$$

By the assumption on $\operatorname{diam}(Q_i)$ we have $M_i \lesssim \|u_0\|_\infty(\Delta x)^d$ for every i, and also by assumption $N \lesssim (\Delta x)^{-d}$, so $NM_i \lesssim 1$ for every i and

$$\begin{aligned}&\mathbb{E}_{x_0}[\mathrm{W}_1(\mathbb{E}_\omega[X_\delta^\nu(t,\cdot)_\# u_0], u_{EM}(t,\bar{x}_0))^2] \qquad (6.3.24)\\ &\lesssim (\Delta x)^{2-2\alpha}\|u_0^i\|_{L^1(\mathbb{T}^d)} + \mathbb{E}_\omega[\mathscr{L}(K_\omega^c)] + \|u_0\|_\infty\int_{\mathbb{T}^d}\mathbb{E}_\omega[|X_{EM}(t,x_0)-X_\delta^\nu(t,x_0)|]dx_0\\ &\lesssim (\Delta x)^{2-2\alpha} + \frac{|\log(\Delta x)|^{-p}}{\alpha^p} + C_{\nu,\Delta t}|\log(\Delta t)|^{-1}.\end{aligned}$$

Estimates (6.3.24) and (6.3.19) let us conclude and get (6.3.9). □

Before giving a probabilistic approximation result based on a Monte Carlo method, like the one given in Proposition 5.3.1, we also need the following result, which we state in a non-sharp version which is sufficient for our needs.

Lemma 6.3.1. ***[79]***
Let μ_N be the empirical distribution associated to the N-sample of a probability distribution μ in $\mathbb{T}^d$. Then :

$$\mathbb{E}_\omega[W_1(\mu,\mu_N)] \le C\frac{1}{\sqrt{N}}$$

and for every $0 < x \le 1$

$$P(\mathrm{W}_1(\mu,\mu_N) > x) \le K\exp(-cNx^2)$$

for some constants $K, c, C > 0$.

Proposition 6.3.3. ***Monte Carlo estimate of the solution of*** (ADE)
Let b, u_0 and $\{Q_i\}_{i=1}^N$ satisfy the assumptions of Theorem 6.3.1. Moreover, assume for simplicity $\alpha = 1/2$, $\Delta t = \Delta x = h$ for some $h > 0$ and that $C_{t,\nu\Delta t}|\log(h)|^{-1} \gg \sqrt{h}$, where $C_{\nu,\Delta t}$ is the constant defined in Proposition 6.3.2. With a slight abuse of notation we denote with $\bar{u}_{EM}(t,\bar{x}_0,\omega)$ the measure obtained considering a single realisation of the Brownian motion and using it in the definition of X_{EM} to advect the Dirac deltas whose starting positions are $\bar{x}_0 = (x_0^1,\dots,x_0^N)$. Then, we can define:

$$S_{n,m}(t) := \frac{1}{n}\sum_{i=1}^{n}\underbrace{\left(\frac{1}{m}\sum_{j=1}^{m}\bar{u}_{EM}(t,\bar{x}_0^i,\omega_j)\right)}_{:=S_m(t,\bar{x}_0^i,\bar{\omega})} = \frac{1}{n}\sum_{i=1}^{n} S_m(t,\bar{x}_0^i,\bar{\omega}),$$

where $\bar{\omega} = (\omega_1,\dots,\omega_m)$. We are thus considering the average of n realisations of $S_m(t,\bar{x}_0,\omega)$ where each S_m is obtained, for each choice of $\bar{x}_0$ fixed, by considering the average of m realisations of $\bar{u}_{EM}(t,\bar{x}_0,\omega)$. We recall that the probability $\bar{P}_{x_0}$ with which we choose the Dirac deltas is uniform over each cell of the mesh $\{Q_i\}_{i=1}^N$, while $\bar{\mathrm{P}} = \underbrace{P\otimes\cdots\otimes P}_{m\ \text{times}}$ is the probability of the space $(\Omega^m,\bar{P})$. Then we can estimate "how close" we can expect $S_{n,m}$ to be to the solution u of (ADE) *in 1-Wasserstein distance:*

$$(\bar{\mathrm{P}}_{x_0}\otimes\bar{P})\left(\left\{|\,\mathrm{W}_1(S_{n,m}(t),u(t,\cdot)) - \mathbb{E}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t),u(t,\cdot))]| > k\sqrt{C\left[\frac{C_{t,\nu\Delta t}|\log(h)|^{-1}}{n} + \frac{1}{nm}\right]}\right\}\right) \le \frac{1}{k^2},$$

where

$$\mathbb{E}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))] \lesssim \frac{C_{t,\nu\Delta t}}{|\log(h)|} + \frac{1}{\sqrt{m}}$$

and $C > 0$ is a positive constant.

Proof. By Lemma 5.3.1:

$$\mathrm{W}_1(S_{n,m}(t), u(t,\cdot)) \leq \frac{1}{n}\sum_{i=1}^{n} \mathrm{W}_1(S_m(t,\bar{x}_0^i,\bar{\omega}), u(t,\cdot)).$$

Since each each $S_m(t,\bar{x}_0^i,\bar{\omega})$ is by construction independent and identically distributed and by triangular inequality for the Wasserstein distance, it holds that

$$\mathbb{E}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))] \leq \mathbb{E}_{x_0,\omega}[\mathrm{W}_1(S_m(t,\bar{x}_0,\bar{\omega}), u(t,\cdot))].$$

and

$$\mathrm{Var}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))] \leq \frac{1}{n}\mathrm{Var}_{x_0,\omega}[\mathrm{W}_1(S_m(t,\bar{x}_0,\bar{\omega}), u(t,\cdot))].$$

Using the same notation of Theorem 6.3.1, we define for any $\bar{x}_0 = (x_0^1 \dots, x_0^N)$:

$$u_{EM}(t,\bar{x}_0) := \sum_{i=1}^{N} M_i \mathbb{E}_\omega[\delta_{X_{EM}(t,x_0^i,\omega)}] = \sum_{i=1}^{N} \mathbb{E}_\omega[X_{EM}(t,\cdot,\omega)_\#(M_i\delta_{x_0^i})].$$

By triangular inequality, for every choice of $\bar{x}_0$ and of $\bar{\omega} = (\omega_1, \dots, \omega_m)$, it holds:

$$\mathrm{W}_1(u(t,\cdot), S_m(t,\bar{x}_0,\bar{\omega})) \leq \underbrace{\mathrm{W}_1(u(t,\cdot), u_{EM}(t,\bar{x}_0))}_{1} + \underbrace{\mathrm{W}_1(u_{EM}(t,\bar{x}_0), S_m(t,\bar{x}_0,\bar{\omega}))}_{2}.$$

Since we want to bound the expected value and the variance of $\mathrm{W}_1(u(t,\cdot), S_m(t,\bar{x}_0,\bar{\omega}))$, it will be sufficient to bound the expected value and the variance of 1 and 2. In particular, we will use that

$$\begin{aligned}
&\mathrm{Var}_{x_0,\omega}[\mathrm{W}_1(u(t,\cdot), S_m(t,\bar{x}_0,\bar{\omega}))] \leq \mathbb{E}_{x_0,\omega}[\mathrm{W}_1(u(t,\cdot), S_m(t,\bar{x}_0,\bar{\omega}))^2] \\
&\leq \mathbb{E}_{x_0,\omega}[|\,\mathrm{W}_1(u(t,\cdot), u_{EM}(t,\bar{x}_0)) + \mathrm{W}_1(u_{EM}(t,\bar{x}_0), S_m(t,\bar{x}_0,\bar{\omega}))|^2] \\
&\leq 2\mathbb{E}_{x_0,\omega}[\mathrm{W}_1(u(t,\cdot), u_{EM}(t,\bar{x}_0))^2] + 2\mathbb{E}_{x_0,\omega}[\mathrm{W}_1(u_{EM}(t,\bar{x}_0^i), S_m(t,\bar{x}_0,\bar{\omega}))^2].
\end{aligned}$$

As for 1, we notice that the measures $u(t,\cdot)$ and $u_{EM}(t,\bar{x}_0)$ do not depend on ω, so we can use the bounds we derived in Theorem 6.3.1. Notice that the estimate on the variance we proved in Theorem 6.3.1 was actually an estimate on the second moment, as we can clearly see from the proof. Both the expected value and the second moment of the term 2 will be estimated uniformly for each choice of $\bar{x}_0$. Indeed, by Lemma 6.3.1 we have that

$$\mathbb{E}_\omega[\mathrm{W}_1(u_{EM}(t,\bar{x}_0), S_m(t,\bar{x}_0,\bar{\omega}))] \lesssim \frac{1}{\sqrt{m}},$$

independent on the choice of $\bar{x}_0$. As for the variance, we use the estimate on the superlevel sets given by Lemma 6.3.1 and the layer cake formula to get

$$\begin{aligned}
\mathbb{E}_\omega[\mathrm{W}_1(u_{EM}(t,\bar{x}_0), S_m(t,\bar{x}_0,\bar{\omega}))^2] &= \int_0^{+\infty} 2xP(\mathrm{W}_1(u_{EM}(t,\bar{x}_0), S_m(t,\bar{x}_0,\bar{\omega})) > x)dx \\
&\leq 2K\int_0^{+\infty} x\exp(-cMx^2)dx = \frac{K}{cm} \approx \frac{1}{m}.
\end{aligned}$$

So, since we have chosen $\alpha = 1/2$ and $\Delta t = \Delta x = h$, by Theorem 6.3.1:

$$\mathbb{E}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))] \le \mathbb{E}_{x_0,\omega}[\mathrm{W}_1(u(t,\cdot), S_m(t,\bar{x}_0,\bar{\omega})] \lesssim C_{t,\nu\Delta t}|\log(h)|^{-1} + \frac{1}{\sqrt{m}}$$

and

$$\mathrm{Var}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))] \le \frac{\mathrm{Var}_{x_0,\omega}[\mathrm{W}_1(u(t,\cdot), S_m(t,\bar{x}_0,\bar{\omega})]}{n} \le C\left[\frac{C_{t,\nu\Delta t}|\log(h)|^{-1}}{n} + \frac{1}{nm}\right]$$

for some $C > 0$. The thesis follows by Chebyshev's inequality, since:

$$\begin{aligned}
&(\bar{\mathrm{P}}_{x_0} \otimes \bar{P})\left(\left\{|\,\mathrm{W}_1(S_{n,m}(t), u(t,\cdot)) - \mathbb{E}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))]| > k\sqrt{C\left[\frac{C_{t,\nu\Delta t}|\log(h)|^{-1}}{n} + \frac{1}{nm}\right]}\right\}\right)\\
&\le (\bar{\mathrm{P}}_{x_0} \otimes \bar{P})\left(\left\{|\,\mathrm{W}_1(S_{n,m}(t), u(t,\cdot)) - \mathbb{E}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))]| > k\sqrt{\mathrm{Var}_{x_0,\omega}[\mathrm{W}_1(S_{n,m}(t), u(t,\cdot))]}\right\}\right)\\
&\le \frac{1}{k^2}.
\end{aligned}$$

□

Remark 6.3.1. ***Numerical methods for* (ADE) *with a Sobolev vector field***
Up to our knowledge, the only other numerical method for approximating solutions of (ADE) *when b has merely Sobolev spatial regularity has been given in [128], where the authors prove stability estimates in logarithmic Wasserstein distance for a finite volume method applied to* (ADE)*, building upon the works [142, 143] for the continuity equation. However, there are many results (see Remark 6.3.2) which focus on approximating stochastic flows with rough velocity fields b.*

Remark 6.3.2. ***Sharper results on stochastic flows & possible improvements***
In recent years, the topic of approximating the solution of (SDE) *when b is not regular has gained more and more interest. Such estimates hold under very mild regularity assumptions on b (e.g. a Ladyzhenskaya-Prodi-Serrin integrability condition) and assuming neither a regularity assumption on b nor a condition on $\mathrm{div}_x(b)$. In this sense, we remark once again that the aim of this chapter was not to obtain the sharpest results, but to illustrate how deterministic techniques can work in essentially the same way when dealing with stochastic flows. This is, however, not really surprising: the biggest simplification we introduced was the assumption of an additive Brownian motion, which has the very nice property of cancelling out when subtracting the solution of the Euler-Maruyama scheme and of the stochastic flow at the same time $t \in [0,T)$. Had it been a multiplicative noise, purely deterministic techniques probably would have failed. There are results that use different techniques, native to the stochastic world, that can tackle such a problem and prove better convergence rates. We refer the reader to [97] and the references therein for more details.*

Chapter 7

Diffusion by randomly forced transport: two models in microhydrodynamics

7.1 Introduction: motivation and background

In this chapter, we focus on a modeling problem in biology that can be summarised as follows:

How does random forcing from immersed microorganisms increase diffusivity in a fluid?

Our approach will change as we move from the realm of pure mathematics to that of physics. This means that, on the one hand, we will refrain from providing fully rigorous statements and proofs; on the other hand, we will consider realistic physical situations which, with the necessary simplifications, we aim to model as accurately as possible. Although the content of this chapter can be seen as a continuation of the work presented in the rest of the thesis, the shift in focus from mathematics to physics motivates presenting it as a self-contained study, allowing interested readers to consult it independently from the rest of the thesis. In the first two sections, we outline the motivations and real-world background that led to the formulation of the problem, providing also a preliminary mathematical analysis showing how a velocity field (which we shall ultimately assume to be produced by microorganisms immersed in a fluid) which changes direction randomly gives rise to a diffusive behaviour in the long-time regime. In Sections 7.3 and 7.4, we model two different stirring mechanisms: in one case, the microorganisms stir the fluid by swimming through it; in the other, they do so by anchoring to the floor and feeding.

Aquatic microorganisms inhabit environments where nutrients are scattered and, at the scales relevant to many such organisms, molecular diffusion is typically too slow to provide them resources at a biologically useful rate. For example, the diffusivity of dissolved organic matter in water is on the order of 10^{-9}m^2/s [65], implying that pure molecular transport can require hours to homogenise across millimeter scales. This discrepancy underscores the central challenge of microbial life: how to overcome the limitations of diffusion in a world at low Reynolds number [1]. Since a passive approach where a microorganism "waits" for nutrients to reach it is doomed to fail, active manipulation of the fluid environment is necessary to make such encounters more likely. Through swimming, ciliary beating or flagellar motion, microorganisms generate flows that both transport food particles to their feeding apparatus and stir the surrounding medium. This process, often called *biogenic mixing* or *biomixing*, effectively enhances the diffusivity of the fluid beyond its molecular baseline, increasing the encounter rate between cells and nutrients [117].

We will focus in particular on two different stirring mechanisms, which are employed for instance by *Volvox* and *Stylonychia* respectively. The *Volvox* is a genus of freshwater green algae that form hollow spherical colonies (containing up to several thousands individual cells) embedded in a gelatinous wall (see Figure 7.1).

[1]We recall that the Reynolds number is a dimensionless quantity which quantifies the relative importance of inertial forces with respect to viscous ones. A high Reynolds number means that inertial forces dominate and the flow is more turbulent, while at low Reynolds number viscous forces dominate and the flow is typically more laminar. We refer the reader to [138]. In this work the author explains (in an informal but very informative way) how motion in a fluid is fundamentally different between organisms moving at high Reynolds number (like a human swimming in water) and organisms at low Reynolds number (like microorganisms).

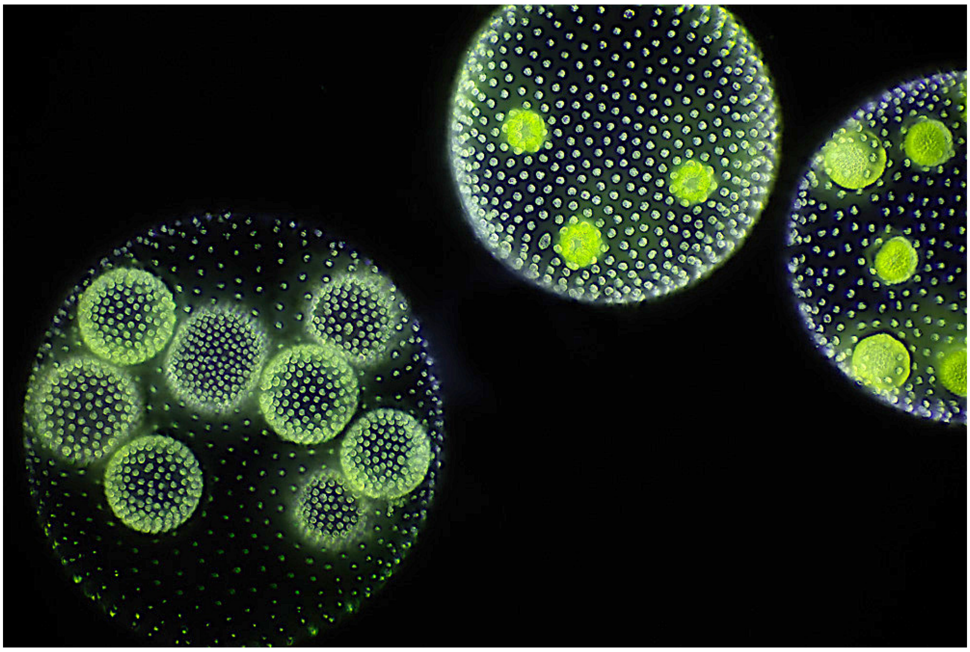
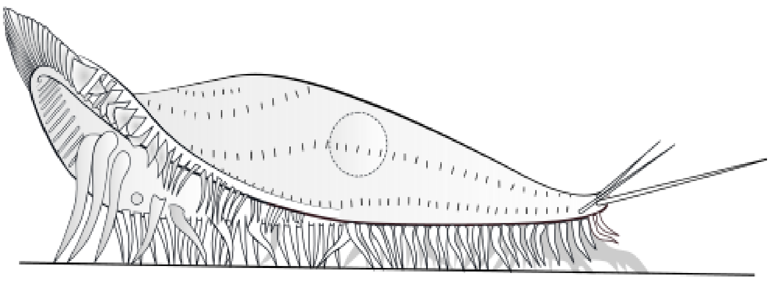


Figure 7.1: Colonies of *Volvox*. Each green dot on the spherical surface is a single cell. Inside each parent colony we can see smaller daughter colonies that develop inside the parent. Image by Frank Fox via Wikimedia Commons, licensed under CC BY-SA 3.0 Germany, now shared under CC BY-SA 4.0.

Each individual cell on the spherical surface has two outward flagella and the combined motion of such flagella allows the whole colony to swim through water: this swimming motion stirs the neighbouring fluid. A *Stylonychia*, instead, is a genus of unicellular ciliates whose feeding mechanism is fundamentally different: they feed by anchoring to the floor and using their adoral zone of membranelles to create a fluid flow which transports nutrients into their feeding apparatus (see Figure 7.2).

(a) Cropped photograph from above of a *Stylonychia Mytilus*. Photograph by Picturepest via Wikimedia Commons, licensed under CC BY 2.0, now shared under CC BY 4.0 .



(b) Part of the drawing of a *Stylonychia Mytilus* walking on its cirri, viewed from one side. Image by Deuterostome via Wikimedia Commons, licensed under CC BY-SA 4.0, now shared under the same license.

Figure 7.2

We will study two stirring mechanisms produced by microorganisms which either swim through water (like *Volvox*) or anchor to the ground and "pump" fluid into their feeding apparatus (like *Stylonychia*). We will use different models for the action of the microorganisms: the *squirmer model* and the *regularised Stokeslet model*. The squirmer model, originally introduced by Lighthill [116] and Blake [22], idealises a swimming microorganism as a sphere with a prescribed tangential velocity on its surface, see also [107] for a more modern presentation. Building on this model, it is shown in [152, 117] that the enhanced diffusion produced by randomly distributed microorganisms depends primarily on far-field stresslet contributions and on random changes of the swimming direction. We will give a more detailed analysis of the works [152, 117] in Section 3. In Section 4, we analyse the other stirring mechanism employed for instance by *Stylonychia*, modeling the microorganisms as regularised Stokeslets [53] placed at some fixed height above a no-slip floor. The presence of the floor greatly impacts the generated flow [23] and it breaks the three-dimensional isotropy that was assumed in [152, 117]. For a detailed experimental analysis of the enhanced diffusion produced by *Stylonychia* [2], we refer the reader to [155].

We conclude this section by noting that the importance of biomixing is not limited to the microscopic world: at ecological scales, biomixing has been conjectured to influence processes such as the vertical transport of nutrients in the ocean. It is argued [68, 99] that the cumulative stirring produced by zooplankton and other aquatic swimmers could contribute significantly to ocean mixing. Although the magnitude of this contribution remains debated [158, 83], there is broad agreement that microorganism-induced flows reshape nutrient landscapes in ways that profoundly affect microbial community dynamics and trophic interactions [150, 151].

7.2 The connection between random transport and diffusion

In this section, we explore how random reorientations of the velocity field driving particles motion ultimately lead to an *enhanced diffusion*, focusing in particular on the three-dimensional case. We recall that if a particle trajectory follows a stochastic differential equation [3] of the form

$$dX_t = b(t, X_t)dt + \sqrt{2\nu}dW_t, \tag{7.2.1}$$

with b a deterministic divergence-free velocity field and W_t a d-dimensional Brownian motion, then the probability density function $p(t,x)$ of its position x satisfies the Fokker-Planck equation

$$\begin{cases} \partial_t p + b \cdot \nabla_x p = \nu \Delta_x p & \text{in } (0,T) \times \mathbb{R}^3 \\ p(0,x) = p_0(x) \end{cases}, \tag{7.2.2}$$

with p_0 an initial probability distribution. We will see how, when b "changes randomly", the transport term in (7.2.2) is substituted (in a long-time limit we will discuss more in detail) by an additional diffusion term $-\kappa\Delta_x p$, and the probability density p can be obtained as the solution of a diffusion equation with a higher diffusion coefficient $D = \kappa + \nu$:

$$\begin{cases} \partial_t p = D\Delta_x p & \text{in } (0,T) \times \mathbb{R}^3 \\ p(0,x) = p_0(x) \end{cases}. \tag{7.2.3}$$

We recall that the symbol $\mathbb{E}_X[\cdot]$ stands for the expected value of a quantity with respect to the probability distribution of a random variable X, which will depend on the context. Namely, if X is a real-valued random variable with probability density function f, then for any measurable function $g : \mathbb{R} \to \mathbb{R}$ we have

$$\mathbb{E}_X[g(X)] = \int_{\mathbb{R}} g(x) f(x) dx.$$

[2] As the reader may have noticed, in [155] the study is performed on a population of *Stylonychia Lemnae*, while we provided images of *Stylonychia Mytilus*. For our purposes, the difference is completely negligible. The two species are almost indistinguishable and they behave in the same way. We will thus refer to such microorganisms simply with their genus *Stylonychia*, without specifying the species.

[3] We refer the reader which is not familiar with stochastic differential equations to [131, 15, 98] for more details, see also the first section of Chapter 6 for a discussion on the connection between (7.2.1) and (7.2.2).

Obvious generalisations to the d-dimensional case can be made, and we will sometimes write only $\mathbb{E}[\cdot]$ when there is no ambiguity about which probability measure the expectation is considered with respect to. We notice that the solution of (7.2.2) can also be interpreted as the evolution of an initial density p_0, and (7.2.2) is also called *advection-diffusion equation*. However, for the moment we prefer to interpret p and p_0 as probability distributions of the position of a particle.

We shall now show how the diffusion coefficient D of (7.2.3) can be deduced by the Lagrangian behaviour of a particle. We define the *mean squared displacement* of a particle at time T as:

$$\mathrm{MSD}(T) := \mathbb{E}[\|X_T - X_0\|^2],$$

and we will show that

$$\mathrm{MSD}(T) = 2dDT \tag{7.2.4}$$

holds in $\mathbb{R}^d$. We start by considering (7.2.1) assuming $b = 0$, so that (7.2.2) and (7.2.3) coincide. To keep this discussion as self-contained as possible, we give the definition of Brownian motion and also a physical interpretation of why the trajectories of a particle in a still fluid can be modeled as a Brownian motion.

Definition 7.2.1. ***Real-valued Brownian motion, [15]***
A real-valued process $W = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (W_t)_{t\geq 0}, \mathrm{P})$ is a standard Brownian motion if:

- $W_0 = 0$ *almost surely;*
- *for every $0 \leq s \leq t$ the random variable $W_t - W_s$ is independent of $\mathcal{F}_s$* [4]*;*
- *for every $0 \leq s \leq t$ the random variable $W_t - W_s$ has law $N(0, t - s)$, with $N(\mu, \sigma^2)$ the law of a Gaussian distribution with mean μ and variance σ^2.*

If the stochastic process W is multiplied by some positive constant C, we say that CW_t is a (real-valued) Brownian motion.

Definition 7.2.2. ***Multidimensional Brownian motion, [15]***
A $\mathbb{R}^d$-valued process $W = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (W_t)_{t\geq 0}, \mathrm{P})$ is a standard multidimensional (also called d-dimensional) Brownian motion if:

- $W_0 = 0$ *almost surely;*
- *for every $0 \leq s \leq t$ the random variable $W_t - W_s$ is independent of $\mathcal{F}_s$;*
- *for every $0 \leq s \leq t$ the random variable $W_t - W_s$ has law $N(0, (t - s)I)$, with $N(\mu, \sigma^2 I)$ the law of a multidimensional Gaussian distribution with mean μ and covariance matrix $\sigma^2 I$.*

If the stochastic process W is multiplied by some positive constant C, we say that CW_t is a multidimensional Brownian motion.

Remark 7.2.1. ***On the construction of a multidimensional Brownian motion***
It is not hard to verify that a standard d-dimensional Brownian motion can be constructed from the real-valued case. Consider the stochastic process

$$W_t := (W_t^1, \dots, W_t^d)$$

where W_t^i are independent real-valued standard Brownian motions. Then, W_t is a standard d-dimensional Brownian motion.

[4] $(\mathcal{F}_t)_{t\geq 0}$ is a *filtration* over the σ*-algebra* $\mathcal{F}$. This is a technical detail and the reader not familiar with stochastic processes may safely ignore it.

The mathematical description in the above definitions, although formally correct, may be a bit obscure for those who are not already familiar with it, and it does not give a good physical motivation of why a particle in a still fluid should move according to the stochastic process W_t we just defined. We will try to provide such a motivation by giving a physical interpretation of the solution $p(t,x)$ of (7.2.3) based on probabilistic arguments first established by none other than Albert Einstein [72]. In [72] it is shown how the small random fluctuations to which particles are subject in a fluid (described for the first time by the botanist Robert Brown [35], hence the name *Brownian motion*) are caused by repeated collisions with fluid molecules [5] and are ultimately the cause of diffusion [6]. We refer the reader to [91] for a more detailed review on the importance and the history of Brownian motion in both physics and mathematics. We now derive (7.2.3) using basic Brownian motion theory.

First, we give a heuristic explanation of why the trajectories of a particle in a still fluid can be modeled as a Brownian motion via its connection with *random walks* (see [20, 48] for a detailed analysis of random walks in biology). Consider a particle placed at the origin of $\mathbb{R}^3$ and assume it is immersed in a fluid with no external forces acting on it. By Newton's first law it would remain at rest, however fluid molecules continuously collide with it, causing it to move. We assume that:

- Collisions occur every dt seconds.
- Each collision displaces the particle by a vector $\vec{r}$. We assume that there is no preferential direction of displacement, i.e. we assume isotropy, so $\mathbb{E}[\vec{r}] = 0$ and $\mathbb{E}[\|\vec{r}\|^2] = 2d\nu\, dt$ for some constant $\nu > 0$ (we will explain later why the assumption on $\mathbb{E}[\|\vec{r}\|^2]$ is physically justified).
- The direction of each displacement is independent of the others.

If we denote by S_N the position of the particle at time T after $N = T/dt \in \mathbb{N}$ collisions, its position at time T will be given by:

$$S_N = \sum_{i=1}^{N} \vec{r}_i.$$

Then we can easily compute that, assuming $M = T'/dt \in \mathbb{N}$ is smaller than N, the following properties hold:

- $S_0 = 0$;
- $\mathbb{E}_{\{\vec{r}_i\}_{i=1}^N}[S_N - S_M] = 0$ and $S_N - S_M$ is independent of S_M;
- The covariance matrix of $S_N - S_M$ is

$$\mathrm{Cov}(S_N - S_M) = (N - M)\mathbb{E}[\|\vec{r}\|^2]I = 2\nu|T - T'|I.$$

Notice that if we had not assumed $\mathbb{E}[\|\vec{r}\|^2]$ to be proportional to dt, the variance of S_N would either diverge or converge to zero as $dt \to 0$, which is clearly not realistic. The last step would be to show that as $dt \to 0$ the random walk S_N converge to a Brownian motion. This can be proved and it is actually a famous result known as Donsker's invariance principle [98]. So, we have that passing to the limit $dt \to 0$ we can assume the particle to move according to a three-dimensional Brownian motion $\sqrt{2\nu}W_t$. Notice that it is now trivial to deduce (7.2.4), since:

$$\mathrm{MSD}(T) = \mathbb{E}\left[\left\|\sqrt{2\nu}W_T\right\|^2\right] = 2d\nu T.$$

[5] At the time of Einstein's article [72] the atomic model was not universally accepted as it is today. The results in [72], later confirmed by the experiments of Jean Baptiste Perrin [136], provided strong support for the atomic model.

[6] It is worth to notice that the law of diffusion was already well-known long before Einstein's work. As early as 1855 Adolf Fick [76] empirically formulated *Fick's laws*. The first of these laws, in particular, states that particles tend to move from regions of high concentration to regions of low concentration, moving directly opposite the concentration gradient. From this observation, it is not difficult to derive (7.2.3). The breakthrough in [72] does not lie in the derivation of (7.2.3), but rather in the description of the microscopic physical mechanisms causing it.

The last thing we need to prove is that the probability density function $p(t,x)$ of the position x of the particle satisfies (7.2.3).

Given $p(t,x)$ the probability density function of the particle at time t, we have that at time $t+dt$ it satisfies:

$$p(t+dt,x) = \int_{\mathbb{R}^d} p(t,x-v)f(v)dv, \tag{7.2.5}$$

where v is the displacement the particle undergoes in the time interval $[t,t+dt]$ and f is the probability density function of the displacements v. Namely, the probability that the particle is at position x at time $t+dt$ is given by the integral over all possible displacements v of the probability that the particle was at $x-v$ at time t, multiplied by the probability that it undergoes a displacement v during the time interval $[t,t+dt]$. Assuming dt is small, so will be v (with high probability at least, since W_{dt} has Gaussian distribution with covariance matrix proportional to dt). By Taylor expanding $p(t,x-v)$ on the right-hand side:

$$p(t,x-v) = p(t,x) - v\cdot\nabla_x p(t,x) + \frac{1}{2}\partial_{x_i}\partial_{x_j}p(t,x)v_iv_j + r(t,x), \tag{7.2.6}$$

with r a term containing higher-order moments. Since we know the law of $v = \sqrt{2\nu}W_{dt}$, it is easy to compute that

$$\int_{\mathbb{R}^d} p(t,x-v)f(v)dv = \mathbb{E}_v[p(t,x-v)] = p(t,x) + \nu\Delta_x p(t,x)dt + o(dt).$$

By Taylor expanding the left-hand side in (7.2.5) we have

$$p(t+dt,x) = p(t,x) + \partial_t p(t,x)dt + o(dt), \tag{7.2.7}$$

thus

$$p(t,x) + \partial_t p(t,x)dt + o(dt) = p(t,x) + \nu\Delta_x p(t,x)dt + o(dt) \implies \partial_t p(t,x) = \nu\Delta_x p(t,x).$$

We now consider (7.2.1) with $b \neq 0$, and we impose the following additional assumptions:

- The velocity field b changes randomly so that the displacements a particle undergoes in the time intervals $[k\tau,(k+1)\tau]$ are independent of each other and of zero mean;
- $\mathbb{E}[\|X_{(k+1)\tau} - X_{k\tau}\|^2] = (\sigma_b^2\tau + 2d\nu)\tau$ for some constant $\sigma_b > 0$, where the term $2d\nu\tau$ comes from the term $\sqrt{2\nu}W_t$ in (7.2.1) .

We can carry out the same analysis as in the previous case, where $b=0$. Since we will be interested in the long-time behaviour of $p(t,x)$, i.e. on $p(T,x)$ for a time $T \gg \tau$, we can rescale the time and assume $T=1$ and $\tau \ll 1$, so

$$p(t+\tau,x) = \int_{\mathbb{R}^d} p(t,x-v)f(v)dv, \text{ with } v = (X_{t+\tau} - X_t) + \sqrt{2\nu}(W_{t+\tau} - W_t).$$

By Taylor expanding the integrand:

$$p(x-v,t) = p(x,t) - v\cdot\nabla_x p(t,x) + \frac{1}{2}v_iv_j\partial_{x_i}\partial_{x_j}p(t,x) + r(t,x), \tag{7.2.8}$$

where r is a remainder term containing higher-order moments and v_i are the components of v. We can easily compute that

- $\mathbb{E}[v_i] = 0$ for every i;
- $\mathbb{E}[v_iv_j] = \left(\frac{\sigma_b^2\tau}{d} + 2\nu\right)\tau\delta_{ij}$, with δ_{ij} a Kronecker delta;
- $\mathbb{E}[v_iv_jv_k] = 0$;

- higher-order moments are $o(\tau)$.

If we Taylor expand the left-hand side we have:

$$p(t+\tau,x) = p(t,x) + \tau\partial_t p(t,x) + o(\tau),$$

which leads to

$$\partial_t p(t,x) = \left(\underbrace{\frac{\sigma_b^2\tau}{2d}}_{=\kappa} + \nu \right) \Delta_x p(t,x) = \underbrace{(\kappa+\nu)}_{=D} \Delta_x p(t,x),$$

so the random reorientation of the velocity field (or, to be more precise, the random displacements it produces every τ seconds) has the effect of an additional diffusion term $\kappa\Delta_x p$ with $\kappa = \frac{\sigma_b^2\tau}{2d}$. Moreover, the relation (7.2.4) between the diffusion coefficient $D = \kappa+\nu$ and the mean squared displacement still holds, since (assuming for simplicity that $T = N\tau$ for some $N \in \mathbb{N}$):

$$\begin{aligned} \mathrm{MSD}(T) &= \mathbb{E}\left[\left\| \left(\sum_{i=0}^{N-1} X_{(i+1)\tau} - X_{i\tau} \right) + \sqrt{2\nu} W_T \right\|^2 \right] = \underbrace{N\tau}_{=T} \sigma_b^2 \tau + 2d\nu T \\ &= 2d \underbrace{\left(\frac{\sigma_b^2\tau}{2d} + \nu \right)}_{=D} T = 2dDT. \end{aligned}$$

7.3 Stirring by swimmers

In this section, we study the enhanced diffusion produced by a specific stirring mechanism, i.e. swimmers moving in random directions. We will describe the model and the results obtained in [117], based on previous results from [152]. We start by describing the simpler model of [152], and later we will use similar techniques to study the more general case considered in [117].

In [152], the authors consider the enhanced diffusivity κ produced by an ensemble of swimmers in an inviscid fluid, whose motion induces a potential flow [106]. The precise setting is the following. Consider M swimmers of characteristic size ℓ in a volume $V \subset \mathbb{R}^3$. We assume that each swimmer moves along a straight line with constant velocity U, with a random swimming direction which is uniformly chosen among all possible orientations. As we saw in section 7.2, in order to calculate the enhanced diffusivity produced by the swimmers it suffices to compute the mean squared displacement (MSD) of a generic fluid particle advected by the vector field the swimmers produce. The main assumption in [152] (which we will always assume, even in the next section) is that the number density $n = M/|V|$ of the swimmers (where $|V|$ denotes the volume of V) is very low, so we can assume that the relevant displacements of a particle are caused by one swimmer at a time (i.e. only when a swimmer is sufficiently near[7] the particle) and at well-separated times.

We now compute the mean squared displacement of a particle (which we will also call *tracer*) initially placed at $x_0 \in V$. Without loss of generality, assuming the volume V is sufficiently large, we can assume $x_0 = (0,0,0) \in \mathbb{R}^3$. We consider a "range of influence" of radius R centered at the particle such that if a swimmer's trajectory does not intersect the ball, we assume its effect on the tracer to be negligible and no displacement occurs. If it does intersect the ball (an event which we will call *impact* from now on), we

[7]The velocity field induced by a sphere moving in an inviscid flow can be explicitly calculated and it decays as r^{-3}, with r the distance from the moving sphere, see for instance [106, Chapter 1, Section 10, Problem 2]. With this rapid decay, we can assume that relevant displacements only occur when a swimmer and the particle are near to each other, which for number densities $n \ll 1$ occurs rarely and typically for one swimmer at a time.

will denote with a the "impact parameter" of the swimmer, namely the minimum distance between the line containing the swimmer's trajectory and the particle. After a time T the tracer will have had N impacts with a swimmer and its final position x_T will be the sum of N displacements, namely

$$x_T = \sum_{i=1}^{N} \Delta(a_i)\hat{r}_i,$$

where $\Delta(a_i)$ is the length of the displacement caused by the i-th impact with impact parameter $a_i \leq R$ and $\hat{r}_i$ is a random unitary vector uniformly distributed among all possible directions. By isotropy, we can assume that the displacements $\Delta(a_i)\hat{r}_i$ are independent from each other. So, the mean squared displacement after a time T is:

$$\mathrm{MSD}(T) = \mathbb{E}_{\bar{a},\bar{\varphi}}[\|x_T\|^2] = \sum_{i=1}^{N} \mathbb{E}_{\bar{a},\bar{\varphi}}[\Delta^2(a_i)] + \sum_{i \neq j} \mathbb{E}_{\bar{a},\bar{\varphi}}\left[\, \Delta(a_i)\Delta(a_j)\hat{r}_i \cdot \hat{r}_j \right],$$

where the expected value $\mathbb{E}_{\bar{a},\bar{\varphi}}[\cdot]$ is considered with respect to possible impact parameters $\bar{a} = (a_1, \ldots, a_N)$ and directions $\bar{\varphi} = (\varphi_1, \ldots, \varphi_N)$ a swimmer can approach the particle from. By isotropy and independence of the displacements directions, if $i \neq j$:

$$\mathbb{E}_{\bar{a},\bar{\varphi}}\left[\, \Delta(a_i)\Delta(a_j)\hat{r}_i \cdot \hat{r}_j \right] = \underbrace{\mathbb{E}_{a_i,\varphi_i}[\Delta(a_i)\hat{r}_i]}_{=0}\,\underbrace{\mathbb{E}_{a_j,\varphi_j}[\Delta(a_j)\hat{r}_j]}_{=0} = 0 \tag{7.3.1}$$

and

$$\sum_{i=1}^{N} \mathbb{E}_{\bar{a},\bar{\varphi}}[\Delta^2(a_i)] = N\mathbb{E}_{a,\varphi}[\Delta^2(a)],$$

so

$$\mathrm{MSD}(T) = N\mathbb{E}_{a,\varphi}[\Delta^2(a)] = N\mathbb{E}_{a}[\Delta^2(a)],$$

where we can consider a fixed φ since the norm of the displacements only depends on the impact parameter a and not on the direction φ the swimmer comes from. We now estimate the number of impacts N. In order to do that, we consider a swimmer initially at distance $L \gg \ell$ from the particle (see Figure 7.3).

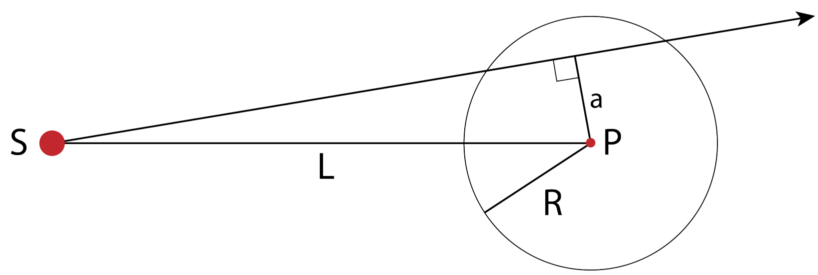


Figure 7.3: A swimmer S coming from a distance L with impact parameter a. The particle is in P with an interaction disc of radius R.

The probability of a swimmer at distance L to collide with the particle (i.e. to intersect the interaction disc of radius R) can be approximated as $R^2\pi/(4L^2\pi) = R^2/(4L^2)$. Since the number density is n, the average number of swimmers in a shell of radius L is $4\pi L^2 nd L$, so the probability that a swimmer at distance L impacts the fluid particle is $\pi R^2 nd L$, and integrating from 0 to UT (we recall that each swimmer's velocity is U) we have that

$$N = \pi R^2 U n T.$$

Thus:

$$\mathrm{MSD}(T) = \pi R^2 U n T \mathbb{E}_a[\Delta^2(a)] = \pi R^2 U n \left(\int_0^R p(a)\Delta^2(a)da \right) T,$$

where $p(a)$ is the probability density function of the impact parameter a. The probability of the impact parameter a being less than x (assuming an impact occurs and the swimmers is moving horizontally from left to right) can be approximated with

$$\mathrm{P}(\{a \leq x\}) = \frac{x^2 \pi}{R^2 \pi} = \frac{x^2}{R^2},$$

hence

$$p(x) = \frac{d}{dx} \mathrm{P}(\{a \leq x\}) = \frac{2x}{R^2}.$$

So we have:

$$\mathrm{MSD}(T) = 2\pi U n T \int_0^R a\Delta^2(a)da,$$

and since $\Delta(a)$ decays very fast in the far field we can send $R \to +\infty$ without any noticeable effect on the integral and get

$$\mathrm{MSD}(T) = 2\pi U n T \int_0^\infty a\Delta^2(a)da. \tag{7.3.2}$$

The diffusivity κ in three-dimensions can be found using the relation (7.2.4), i.e.

$$\mathrm{MSD}(T) = 6\kappa T.$$

By making (7.3.2) non-dimensional, e.g. changing variables with $r = a/\ell$ and dividing and multiplying $\Delta^2(a)$ by ℓ^2, we find

$$\mathrm{MSD}(T) = \left(2\pi \int_0^\infty r \frac{\Delta^2(\ell r)}{\ell^2} dr \right) U n \ell^4 T.$$

In the case of spherical swimmers, it is computed in [152] that $\kappa = 0.266 U n \ell^4$. Moreover, the dominant contribution comes from head-on collisions, since:

$$\int_0^\infty a\Delta^2(a)da \approx 0.254\ell^4 \text{ and } \int_0^\ell a\Delta^2(a)da \approx 0.250\ell^4.$$

For more details (like the two-dimensional case or the case of cylindrical swimmers) we refer the reader to [152]. In [117], the previous analysis is generalised to viscous fluids with M swimmers which move at constant speed U for a fixed length λ, which we assume to be much bigger than the characteristic length ℓ of the swimmers [8], before stopping and changing the direction of swimming. In this viscous case the swimmers are modeled as *squirmers*, i.e. spherical bodies with a prescribed velocity field on the boundary that allows them to swim. The squirmer model is also used for microorganisms whose swimming apparatus is not a collection of cilia on their body, but elongated flagella. There are two main mechanisms with which a squirmer can move, that can summarised as *pushing* the fluid or *pulling* it (see Figure 7.4). Since we are assuming that no external forces are present and that the swimmers move in free space, the net force on the fluid must be zero and in the far field the squirmers can be represented as force dipoles (*stresslets* [107] to be precise) with a given strength β. The parameter β (positive for pullers and negative for pushers) will

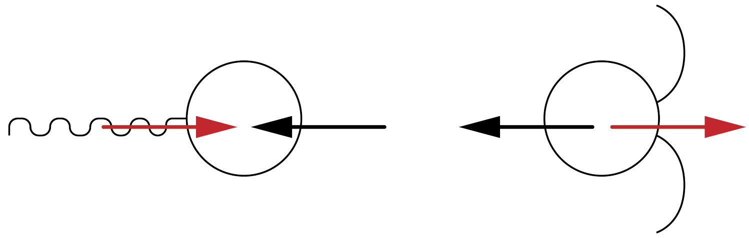


Figure 7.4: A pusher (on the left) and a puller (on the right) swimming by moving their flagella. In red we represent the thrust force and in black the drag of the fluid. Since they are equal in magnitude but opposite in direction, they give rise to a force dipole. Examples of pullers are the algae *Volvox* and *Chlamydomonas Reinhardtii*; examples of pushers are *Escherichia Coli* and sperm cells.

be crucial in estimating the enhanced diffusion produced by the squirmers. For a precise derivation of the squirmer model and more details we refer the reader to [116, 22, 107].

As in the previous case we consider an interaction disc of radius R centered in a fluid particle initially positioned at $x_0 = (0,0,0) \in \mathbb{R}^3$. The position x_T of the particle after a time T is given by the sum of N displacements Δ_λ, each caused by its impact with a single squirmer (we still assume a low number density) which crosses the interaction disc. So:

$$x_T = \sum_{i=1}^{N} \Delta_\lambda(a_i, b_i)\hat{r}_i,$$

where a_i is an impact parameter defined as in [152]: it is the minimum distance between the swimmer's trajectory and the particle position. The parameter b_i is the distance (with sign) between the starting position of the swimmer's trajectory and its point of closest approach to the fluid particle, while $\hat{r}_i$ are independent uniformly distributed unitary vectors (see Figure 7.5). We need to compute the mean squared displacement of the particle, averaging over all possible impact parameters a, b and directions of approach φ. As in the previous case, by isotropy and independence of the displacements the mixed term are zero, i.e. it holds:

$$\begin{aligned} \mathrm{MSD}(T) &= N\mathbb{E}_{a,b,\varphi}[\Delta_\lambda^2(a,b)] + \underbrace{\sum_{i\neq j} \mathbb{E}_{a,b,\varphi}[\Delta_\lambda(a_i,b_i)\Delta_\lambda(a_j,b_j)\hat{r}_i\hat{r}_j]}_{=0} \\ &= N\mathbb{E}_{a,b}[\Delta_\lambda^2(a,b)], \end{aligned}$$

Where in the last term we do not average over φ since the norm of the displacements depends only on a, b and not on the direction of approach φ, which we can assume to be horizontal from left to right, as in Figure 7.5. In order to compute the average number of impacts N in a time T, we assume for simplicity that all swimmers change directions at the same time. That is, at time 0 they all start moving with the same velocity U and once they have traveled a distance λ they stop, change direction and start swimming again. The time it takes the swimmers to travel a distance λ is λ/U, and in that time the probability that there is an impact is

$$\frac{M}{|V|}\int_\Lambda dx = \frac{M|\Lambda|}{|V|}, \tag{7.3.3}$$

[8]This assumption may seem arbitrary, but it is actually *very* physical. Quoting verbatim from [138]:"At low Reynolds number you can't shake off your environment. If you move, you take it along; it only gradually falls behind". That is, if a microorganism wants to explore "new" parts of the fluid looking for more food, it has to travel a relatively long distance.

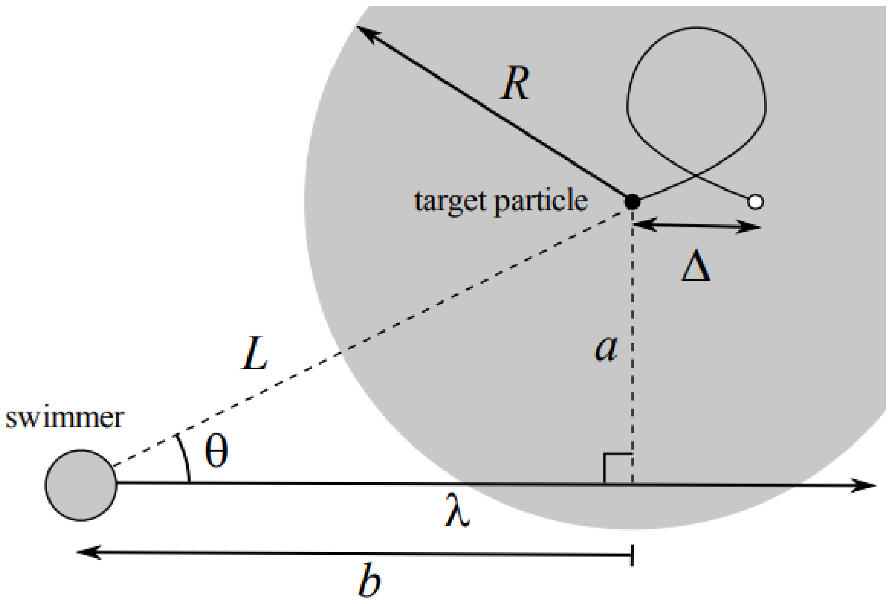


Figure 7.5: Impact parameters a and b, displacement $\Delta = \Delta_\lambda(a,b)$ and swimming path length λ. Notice that when b is positive the swimmer has yet to pass the point of closest approach to the particle, while for negative b it starts its motion past that point. The interaction disc of radius R is also shown. Image from [117], reproduced with permission; it is not included under the CC BY 4.0 license of this thesis.

where $\Lambda \subset V$ is the volume of the fluid "explored" by a single swimmer during a flight of length λ. Notice that an impact occurs if and only if the tracer belongs to one of the volumes Λ explored by the swimmers (see Figure 7.6), whose collective volume we assume to be $M|\Lambda|$, since intersections are unlikely under the hypothesis of low number density.

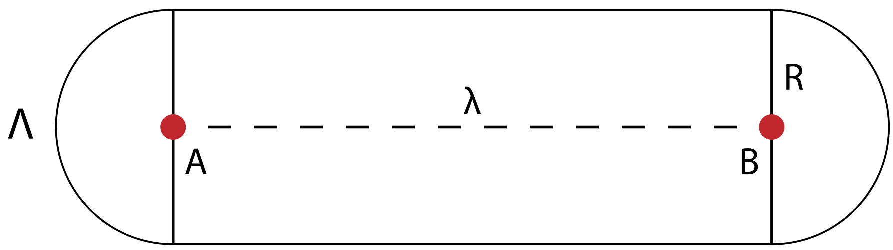


Figure 7.6: An example of volume Λ, which consists of a cylinder of height λ with circular base of radius R, with a half-sphere of radius R attached to each end. The swimmer starts its path from A and it arrives to B. It impacts the particle (assuming an interaction disc of radius R) if and only if the particle lies inside Λ.

We can rewrite the volume integral in (7.3.3) as a function of the impact parameters a, b and since for each swimmer there are UT/λ flights of length λ in a time T, the average number of impacts is

$$\frac{UT}{\lambda}\frac{M}{|V|}\int_0^R \int_{-\sqrt{R^2-a^2}}^{\lambda+\sqrt{R^2-a^2}} 2\pi a \, db \, da = \frac{2\pi n UT}{\lambda}\int_0^R \int_{-\sqrt{R^2-a^2}}^{\lambda+\sqrt{R^2-a^2}} a \, db \, da.$$

We can estimate the mean squared displacement in a very similar fashion. Indeed, we computed N as

$$N = \frac{M}{|V|}\frac{UT}{\lambda}\mathbb{E}_{a,b}[\mathbb{1}_\Lambda],$$

i.e. "counting" the number of impacts between the tracer and a swimmer. We can compute $\mathrm{MSD}(T)$ not by counting the average number of impacts, but rather the mean squared displacement per impact, namely:

$$\mathrm{MSD}(T) = \frac{M}{|V|}\frac{UT}{\lambda}\mathbb{E}_{a,b}[\mathbb{1}_\Lambda \Delta_\lambda^2(a,b)] = \frac{2\pi n UT}{\lambda}\int_0^R \int_{-\sqrt{R^2-a^2}}^{\lambda+\sqrt{R^2-a^2}} a\Delta_\lambda^2(a,b)\, db \, da.$$

Since $\Delta_\lambda(a, b)$ decays rapidly for large values of a, b (the velocity field induced by a squirmer with $\beta \neq 0$ decays as r^{-2}, with r the distance from the squirmer) we can assume $R \to +\infty$, and we obtain

$$\mathrm{MSD}(T) = \frac{2\pi n U T}{\lambda} \int_0^\infty \int_{-\infty}^{+\infty} a\Delta_\lambda^2(a, b)\, db\, da. \tag{7.3.4}$$

We can rescale (7.3.4) to have a non-dimensional integrand and derive scaling laws. If we use the variables $\log(a/\ell)$ and b/λ and we divide and multiply by ℓ^4 the squared norm of the displacement:

$$\mathrm{MSD}(T) = \left(2\pi \int_{\mathbb{R}^2} \frac{a^2\Delta_\lambda^2(a, b)}{\ell^4}\, d\log(a/\ell)\, d(b/\lambda) \right) U n \ell^4 T. \tag{7.3.5}$$

In Figure 7.7, we show the plot of the integrand in (7.3.5) for both the potential flow case and the viscous case with fixed parameters.

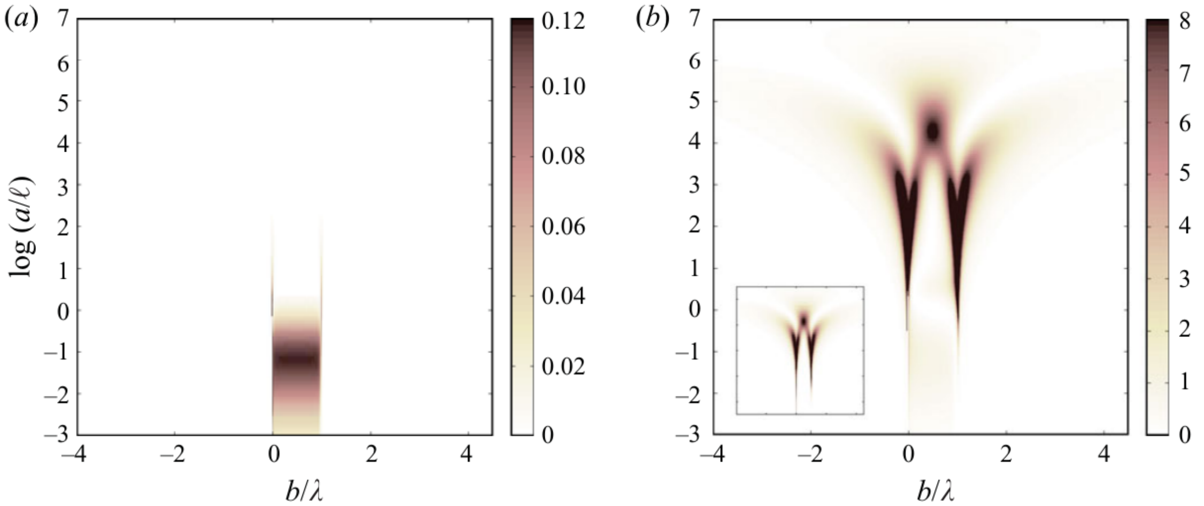


Figure 7.7: Integrand $\ell^{-4}a^2\Delta_\lambda^2(a, b)$ from (7.3.5) due to (a) a swimmer in potential flow and (b) a squirmer with $\beta = 5$. It is clear how the far-field contribution is dominant in the second case, with head-on collisions being barely visible in the plot. The inset on figure (b) represents the integrand corresponding to only the far-field contribution of the flow, i.e. considering the squirmer as a pointwise stresslet of strength β. The computations are performed assuming $U = \ell = 1$ and $\lambda = 100$. Image from [117], reproduced with permission; it is not included under the CC BY 4.0 license of this thesis.

It can be observed that, while in the potential flow case most of the contribution to the integral comes from head-on collisions, in the viscous case it actually comes from the far field. Moreover, in the viscous case the integrand is almost two orders of magnitude larger and an additional phenomenon, absent in the potential flow case, further increases the diffusivity. This phenomenon is known as *Darwin drift* [60, 71, 160] and it consists in the formation of closed regions behind the swimmer (for $\beta > 0$; otherwise they form in front), where vortices capture a particle and drag it along the entire displacement length λ [9] (see Figure 7.8). Finally, in [117] it is also found that for large stresslet parameters β the enhanced diffusivity is more and more dominated by far-field effects, and it behaves like

$$\kappa \approx 2.1\beta^2 U n \ell^4 \tag{7.3.6}$$

(see Figure 7.9), while for $\beta \to 0$ we recover the potential flow case, with $\kappa \approx 0.266 U n \ell^4$.

[9] This is an oversimplification, as noted in [117] in practice there are boundary phenomena and turbolence which make particles enter and exit the wake.

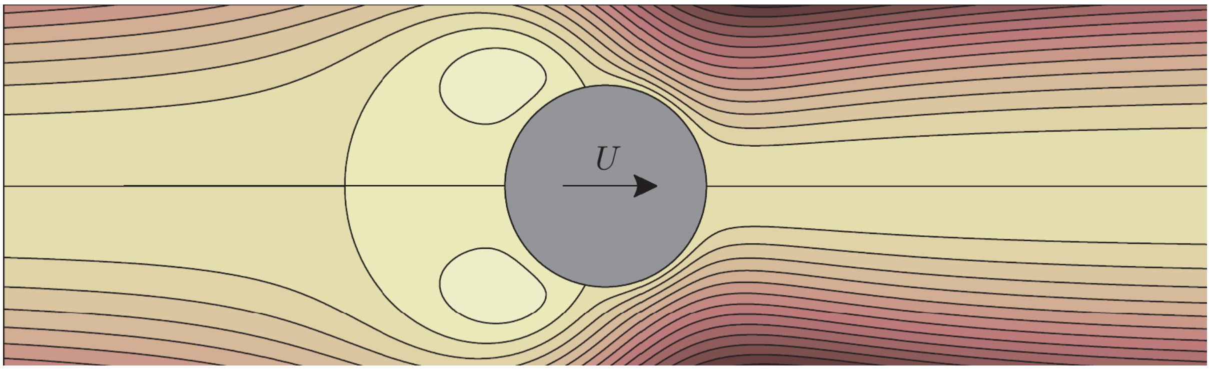


Figure 7.8: Streamfunction contours for a pusher with $\beta = 5$ in a comoving reference frame. We can observe the formation of a closed bubble behind the squirmer where particles remain trapped. Image from [117], reproduced with permission; it is not included under the CC BY 4.0 license of this thesis.

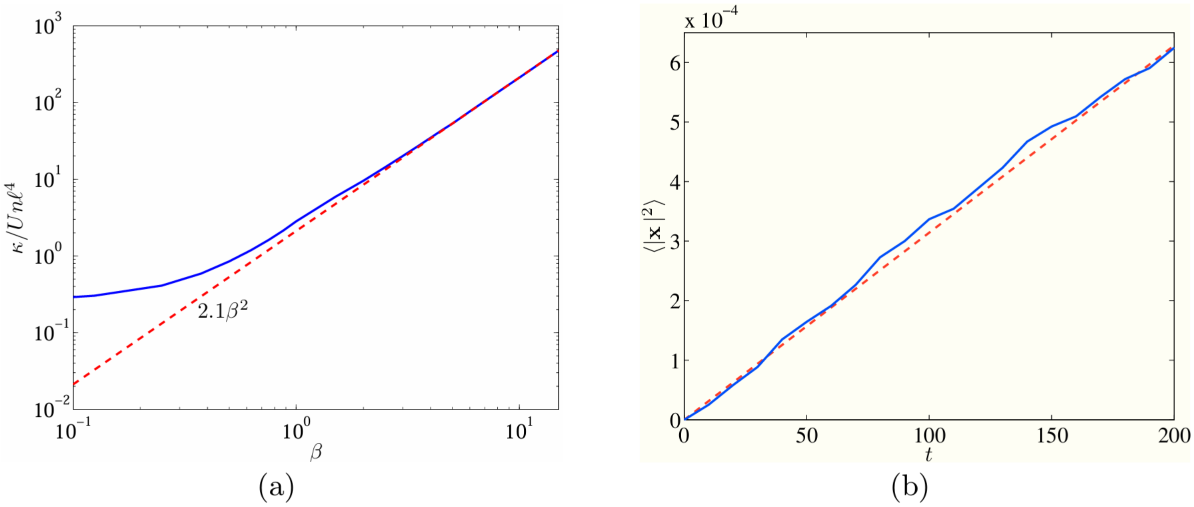


Figure 7.9: (a) The solid line is the enhanced diffusivity κ computed by numerical integration of (7.3.4). The dashed line is the enhanced diffusivity due only to the stresslet term. (b) The solid line is the mean squared displacement computed with 2×10^6 realisations of 10 squirmers with $\ell = U = 1$, $\lambda = 100$ and $\beta = 5$. The dashed line is the mean squared displacement predicted by (7.3.6) with a number density $n = 10^{-8}$. Image from [117], reproduced with permission; it is not included under the CC BY 4.0 license of this thesis.

7.4 Stirring by Stokeslets

7.4.1 Fixed encounter time t_p

We now consider a different stirring mechanism. We assume that an ensemble of microorganisms (which from now on we will call *actuators*) lives on a flat surface. The way such actuators stir the fluid is fundamentally different from the one described in the previous section. While previously the stirring was caused by a swimming motion of zero net force, here we consider microorganisms that crawl on the surface, then stop and anchor themselves to the floor to "pump" fluid into their feeding apparatus. The action of an actuator creates a non-zero net force which can be modeled as a regularised Stokeslet [53]. We recall that a Stokeslet is a fundamental solution of the Stokes equation[10] [100]:

$$\begin{cases} -\nabla_x p + \mu \Delta_x b + F\delta(x) = 0 & \text{in } \mathbb{R}^3 \\ \nabla_x \cdot b = 0 \end{cases}, \tag{7.4.1}$$

[10]The motion of incompressible fluids is, in general, described by the incompressible Navier-Stokes equations. In our case we can neglect the advection term $(b \cdot \nabla)b$ and $\partial_t b$ because both the Reynolds number and the oscillatory Reynolds number are small. See [107, Chapter 2] for more details.

where $\delta(x)$ is a Dirac delta centered at $(0,0,0) \in \mathbb{R}^3$, $F \in \mathbb{R}^3$ is a force vector and μ is the dynamic viscosity of the fluid. The solutions b, p (which are the induced velocity field and pressure in the fluid, respectively) are

$$b(\vec{r}) = \frac{1}{8\pi\mu}\left(\frac{I}{|\vec{r}|} + \frac{\vec{r}\otimes\vec{r}}{|\vec{r}|^3}\right)\cdot F \quad \text{and} \quad p(\vec{r}) = \frac{F\cdot\vec{r}}{4\pi|\vec{r}|^3}, \tag{7.4.2}$$

where $\vec{r}$ is the position vector, $(\vec{r}\otimes\vec{r})\cdot F = (F\cdot\vec{r})\vec{r}$ and I is the identity matrix. A regularised Stokeslet [53] is the solution of the same equation where instead of a Dirac delta (which is numerically hard to treat) we use a smooth function φ_δ supported on a ball of radius δ to approximate the Dirac delta in distribution (a so-called "mollifier" in mathematics). We will consider the parameter δ to be the *radius* of the regularised Stokeslet. In fact, we will not use the velocity field in (7.4.2): the presence of a floor (where we impose a no-slip condition) drastically changes the solution of (7.4.1) both for a Stokeslet and for a regularised Stokeslet placed at height h above the floor. We refer the reader to [23] and [2] for the details and explicit formulas. We just point out that a fundamental difference between a Stokeslet in free space and a Stokeslet whose force is parallel to a no-slip floor is the different decay of the velocity fields they generate. If d is the distance from a Stokeslet, the magnitude of the velocity field in $\mathbb{R}^3$ decays like $1/d$ in the far field, while in presence of a no-slip floor it decays like $1/d^2$. Similarly, a squirmer in $\mathbb{R}^3$ generates a velocity field whose magnitude decays like $1/d^2$ in the far field, while in presence of a no-slip floor it decays like $1/d^3$ [23, 24].

We now describe more in detail the situation we aim to model, together with our main assumptions and objectives. We assume that each actuator can be in two possible states: either active or inactive. When it is active, it is anchored to the floor and it acts as a regularised Stokeslet exerting a force F on the fluid; when not active, it crawls on the floor and we neglect the flow it generates on the fluid [11]. We assume that at any moment the number of *active* actuators (from now on, when considering actuators, we will always implicitly assume they are active) is a given number M and we consider a large surface A on the floor such that the number density of actuators is always $n = M/|A \ll 1$, with $|A|$ the area of A. Such actuators share common features, namely:

- The strength $f = |F|/\mu$ of the Stokeslet they produce, with $|F|$ the modulus of the force acting on the fluid and μ the dynamic viscosity;
- the characteristic size ℓ of the actuators, the height h at which they exert the force F on the fluid, which we assume to be parallel to the ground[12], and the radius δ of the regularised Stokeslet.

Remark 7.4.1. ***Force and strength of a Stokeslet***
We refer to f as the "strength" of a Stokeslet, namely the quantity $f = |F|/\mu$, where $|F|$ is the modulus of the force acting on the fluid and μ is the dynamic viscosity. Instead, when considering F we will call it "force".

Remark 7.4.2. ***Dependence of the displacement on ft_p***
Since the velocity field generated by a Stokeslet (both singular and regularised and both in free space and in presence of a no-slip floor) is the solution of a time-independent PDE which is linear in $f = |F|/\mu$, it is not difficult to verify that the displacement after a time t_p depends only on the product ft_p, assuming a given direction for the force F. That is, a particle is displaced equally by Stokeslets of strength f/c acting for a time ct_p for every positive constant $c > 0$.

We will also make the following assumptions:

- The carpet of actuators refreshes completely every t_p seconds. That is, at time 0 there are M active actuators which stay active up to time t_p, when they all stop pumping and a completely new carpet activates. The positions of the actuators are independent from each other and uniformly distributed;

[11] The specific microorganisms we considered when developing our theoretical model, i.e. *Stylonychia*, do not stop feeding even while moving. However, the force they generate while feeding and the hydrodynamic drag they experience while moving are approximately of the same magnitude but opposite in direction. Hence, they act as a moving stresslet, whose overall effect on the fluid is much less impactful than the one produced by a stationary Stokeslet [155].

[12] The assumption that the force F is parallel to the ground is not really necessary and we only assume it for simplicity. All the arguments carry over to the case of forces with non-zero z components if their component in the xy-plane is uniformly distributed.

- the tracer particle always remains near the floor, and we assume that only the nearest actuator (which we will also call *nearest neighbour*) affects it [13]. When a tracer is advected by the action of its nearest neighbour, we say that the tracer and the actuator have an *encounter*;
- the nearest neighbour remains the same until the carpet refreshes. In other words, the actuators are typically spread far enough so that the displacement produced by a Stokeslet of strength f acting for a time t_p is unlikely to bring the tracer particle closer to a different actuator than the initially nearest one. This hypothesis will be relaxed in the next subsection, where we allow the encounter duration between a tracer and an actuator to be random, whereas under the current assumptions we have that an encounter always lasts for a time t_p;
- the direction of the force F, which we recall we assume to be planar, is uniformly distributed.

As in the previous section, we will work assuming a small number density n of active actuators, but there is an additional difficulty to overcome. Namely, while in the squirmer's case the problem was completely isotropic, now there is a dependence on the height of the tracer: if a particle starts at height z_1, after a given time T its height will have changed randomly, and the action of a Stokeslet at different heights varies greatly. For this reason we will not compute the three-dimensional mean squared displacement, but rather the following quantities:

- $\mathrm{MSD}_{xy}(T, z_1)$, i.e. the mean squared displacement, with respect to the x and y coordinates only, of a tracer initially placed at $(0, 0, z_1)$ after a time T;
- $p_z(z, T, z_1)$, i.e. the probability distribution, after a time T, of the height z of a tracer initially placed at $(0, 0, z_1)$.

Once we have derived $p_z(z, T, z_1)$ we can also compute the mean squared displacement in the z-direction $\mathrm{MSD}_z(T, z_1)$ as

$$\mathrm{MSD}_z(T, z_1) = \int_0^{+\infty} (z - z_1)^2 p_z(z, T, z_1) dz.$$

It is not difficult to see that the total mean squared displacement is simply the sum of the two components:

$$\mathrm{MSD}(T, z_1) = \mathrm{MSD}_{xy}(T, z_1) + \mathrm{MSD}_z(T, z_1)^{14}.$$

We consider a particle starting at position $x_0 = (0, 0, z_1)$, for some $z_1 > 0$. Under the above assumptions, in a time $T = Nt_p$ there are N encounters between a tracer and an actuator: contrary to the approach used in [152, 117] for squirmers, there are no fictitious interaction discs involved. Rather, as we already remarked, we assume that at each activation of the carpet of actuators the tracer is under the effect of only its nearest neighbour, an assumption which we recall is only feasible if n is small. Denoting (with a little abuse of notation) as $x_{xy}(T)$ the xy coordinates of the tracer at time T, we have:

$$x_{xy}(T) = \sum_{i=1}^{N} \Delta_{xy}(r_i, \varphi_i, \theta_i, z_i) = \sum_{i=1}^{N} \Delta^i_{xy},$$

where $\Delta_{xy}(r, \varphi, \theta, z)$ is the displacement in the xy components of a particle at height z caused by the action of an actuator whose polar coordinates, relative to the projection on the plane $\{z = 0\}$ of the particle, are (r, φ) and which is acting on the fluid with a horizontal force F in direction θ. See Figure 7.10 for a visual representation. Notice that, contrary to the previous section where Δ and Δ_λ denoted the norm of the displacement, here we consider $\Delta_{xy}(r_i, \varphi_i, \theta_i, z_i) = \Delta^i_{xy}$ to be the displacement produced by the i-th encounter. Moreover, we consider z_i as a random variable which depends on the values of r_k, φ_k, θ_k for every $k < i$ and on z_1. Namely, the height of the particle at the i-th encounter depends on its starting height z_1 and on the sequence of *encounter parameters* $\{(r_k, \varphi_k, \theta_k)\}_{k=1}^{i-1}$.

[13] If the particle were far from the floor a far-field approach, where the collective action of all actuators is considered, would be more suited. This situation has been investigated in [89].

[14] From now on we will often omit explicitly writing the dependence on the initial height if it is not necessary to specify it.

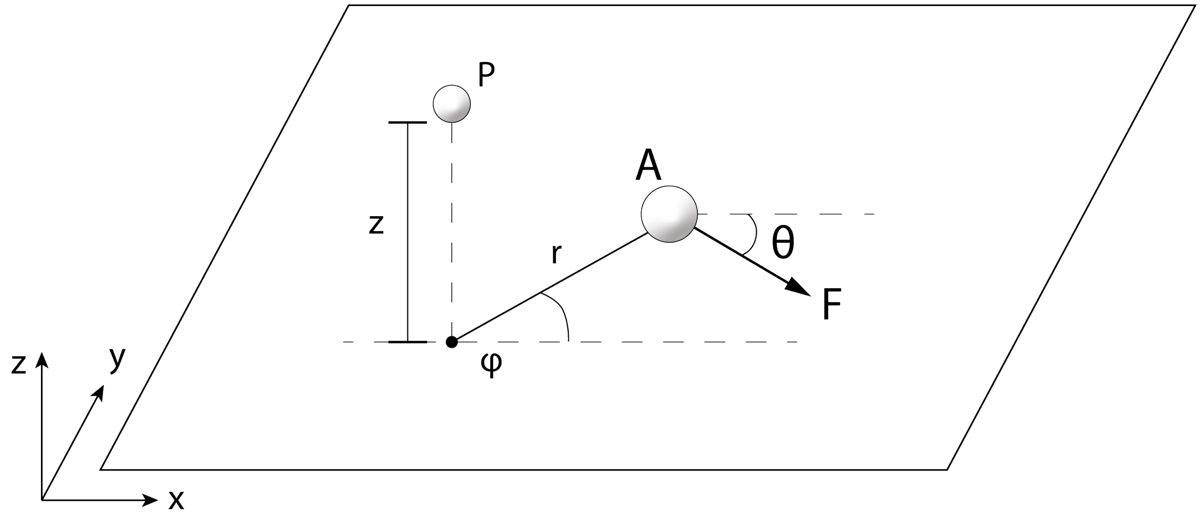


Figure 7.10: Example of an encounter with parameters (r, φ, θ, z). We can see a particle P at height z which is advected by an actuator A with relative position (in polar coordinates) (r, φ) that acts on the fluid with an horizontal force F, which forms an angle θ with the x-axis. Illustration courtesy of Anastasiia Kovalskaia.

Averaging over all possible encounter parameters $\bar{r} = (r_1, \dots, r_N), \bar{\theta} = (\theta_1, \dots, \theta_N)$ and $\bar{\varphi} = (\varphi_1, \dots, \varphi_N)$:

$$\mathrm{MSD}_{xy}(T) = \mathbb{E}_{\bar{r},\bar{\varphi},\bar{\theta}}[\|x_{xy}(T)\|^2] = \sum_{i=1}^{N} \mathbb{E}_{\bar{r},\bar{\varphi},\bar{\theta}}[\|\Delta^i_{xy}\|^2] + \sum_{i\neq j} \mathbb{E}_{\bar{r},\bar{\varphi},\bar{\theta}}[\Delta^i_{xy} \cdot \Delta^j_{xy}]. \tag{7.4.3}$$

We show how the second sum is equal to zero. In the previous section we could use isotropy and independence of displacements, while in this case we already noticed how the presence of a floor breaks isotropy and independence: the displacement a particle undergoes during the second encounter with an actuator depends on the first encounter since it modifies its height. So, while it is easy to see (since isotropy in the xy-directions is maintained) that for every i

$$\mathbb{E}_{\bar{r},\bar{\varphi},\bar{\theta}}[\Delta^i_{xy}] = 0,$$

we cannot conclude the same for the scalar product $\mathbb{E}_{\bar{r},\bar{\varphi},\bar{\theta}}[\Delta^i_{xy} \cdot \Delta^j_{xy}]$ by independence. However, the sum of mixed terms is still zero. Indeed, without loss of generality, assume that $j > i$. We consider a sequence of encounter parameters $\{(r_k, \varphi_k, \theta_k)\}_{k=1}^{j}$ which produce, at the i-th and j-th encounters, horizontal displacements Δ^i_{xy} and Δ^j_{xy} respectively. Now consider a different sequence $\{(\tilde{r}_k, \tilde{\varphi}_k, \tilde{\theta}_k)\}_{k=1}^{j}$ which differs from the previous one only in the last entry, namely:

$$(\tilde{r}_k, \tilde{\varphi}_k, \tilde{\theta}_k) = (r_k, \varphi_k, \theta_k) \quad \text{for every } k = 1, \dots, j-1$$

and

$$(\tilde{r}_j, \tilde{\varphi}_j, \tilde{\theta}_j) = (r_j, \varphi + \pi, \theta_j + \pi).$$

Since the angles φ, θ are uniformly distributed, the two sequences are equally likely; however, in the second case, the j-th displacement will be $\Delta_{xy}(\tilde{r}_j, \tilde{\varphi}_j, \tilde{\theta}_j) = -\Delta^j_{xy}$. Thus, since we are integrating an odd function:

$$\mathbb{E}_{\bar{r},\bar{\varphi},\bar{\theta}}[\Delta^i_{xy} \cdot \Delta^j_{xy}] = 0 \quad \text{for every } i \neq j$$

and

$$\mathrm{MSD}_{xy}(T) = \sum_{i=1}^{N} \mathbb{E}_{\bar{r},\bar{\varphi},\bar{\theta}}[\|\Delta^i_{xy}\|^2].$$

Since the angles θ, φ are uniformly distributed by assumption, we only have to derive the probability density functions of r_i and z_i. We start from the former. We recall that by assumption the only actuator acting on the particle is its nearest neighbour, so we have to compute the probability density function of the minimum distance d of the projection of the tracer on $\{z = 0\}$ from M actuators uniformly distributed over an area A. The probability of d being greater than a given $r > 0$ is the probability that, independently selecting M points uniformly distributed in A, each of them lies outside a circle of radius r centered in the tracer's projection. Thus:

$$\mathrm{P}(\{d > r\}) = \left(\frac{|A| - r^2\pi}{|A|}\right)^M = \left(1 - \frac{r^2\pi}{|A|}\right)^M \approx \exp(-nr^2\pi).$$

We can approximate the probability density function p_r with

$$p_r(r) = \frac{d}{dr}\mathrm{P}(\{d \le r\}) = \frac{d}{dr}[1 - \mathrm{P}(\{d > r\})] = 2\pi rn\exp(-nr^2\pi). \tag{7.4.4}$$

Before deriving the probability density functions of the variables z_i for $i = 1, \dots, N$, we notice how the computation can be simplified by avoiding averaging over φ. Indeed, we will see that there are two relevant quantities for our analysis: one is of course $\|\Delta_{xy}(r, \varphi, \theta, z)\|^2$ and the other one will be the vertical displacement $\Delta_z(r, \varphi, \theta, z)$. We notice that both of them depend only on the difference $\theta - \varphi$ rather than on the individual values of θ and φ. That is, for every z:

$$\|\Delta_{xy}(r, \varphi, \theta, z)\|^2 = \|\Delta^2_{xy}(r, 0, \theta - \varphi, z)\|^2 \quad \text{and} \quad \Delta_z(r, \varphi, \theta, z) = \Delta_z(r, 0, \theta - \varphi, z),$$

so

$$\begin{aligned}
\mathbb{E}_{r,\varphi,\theta}[\|\Delta_{xy}(r, \varphi, \theta, z)\|^2] &= \fint_0^{2\pi}\left(\int_0^\infty \fint_0^{2\pi} \|\Delta^2_{xy}(r, \varphi, \theta, z)\|^2 p_r(r) d\theta dr\right) d\varphi \\
&= \fint_0^{2\pi}\left(\int_0^\infty \fint_0^{2\pi} \|\Delta_{xy}(r, 0, \theta - \varphi, z)\|^2 p_r(r) d\theta dr\right) d\varphi \\
&= \fint_0^{2\pi}\left(\int_0^\infty \fint_0^{2\pi} \|\Delta_{xy}(r, 0, \theta, z)\|^2 p_r(r) d\theta dr\right) d\varphi \\
&= \int_0^\infty \int_0^{2\pi} \|\Delta_{xy}(r, \theta, z)\|^2 \frac{p_r(r)}{2\pi} d\theta dr = \mathbb{E}_{r,\theta}[\|\Delta_{xy}(r, \theta, z)\|^2],
\end{aligned}$$

where from now on we will omit writing the dependence of Δ_{xy} on φ since, for our purposes, we can always assume $\varphi = 0$ [15].

We now find the probability density functions of the heights z_i. We notice that, contrary to other impact parameters which are identically distributed for every encounter, the probability distribution of the height z changes with time, i.e. each z_i has a different distribution $p^i_z(z)$. The key point is to notice that $\{z_i\}_{i=1}^N$ can be considered a discrete-time Markov chain with an uncountably infinite state space [139]. Namely, the probability distribution of z_{i+1} only depends on the value of z_i, and not on the ones before. In general, given a probability distribution of heights $p^i_z(z)$ for the i-th encounter (which in the case of the first encounter is simply a Dirac delta centered in z_1) the following holds:

- The mean squared displacement (in the xy variables) of the i-th encounter can be computed as

$$\begin{aligned}
\mathbb{E}_{r,\theta,z_i}[\|\Delta_{xy}(r, \theta, z_i)\|^2] &= \int_0^\infty p^i_z(z)\left(\int_0^\infty \int_0^{2\pi} \|\Delta_{xy}(r, \theta, z)\|^2 \frac{p_r(r)}{2\pi} d\theta dr\right) dz \\
&= \int_0^\infty p^i_z(z)\mathbb{E}_{r,\theta}[\|\Delta_{xy}(r, \theta, z)\|^2] dz;
\end{aligned}$$

[15] In light of this discussion, it can be easily verified that the assumption of θ being uniformly distributed is not really necessary. Assuming that the actuators are uniformly distributed, the same holds for φ, and it is sufficient that for every φ the distribution $p_{\theta-\varphi}$ of the difference $\theta - \varphi$ is the same (not necessarily uniform). With this assumption, it still holds that $\mathbb{E}[\Delta^i_{xy}] = 0$ and $\mathbb{E}[\Delta^i_{xy} \cdot \Delta^j_{xy}] = 0$ if $i \neq j$, and as we just saw $\|\Delta_{xy}\|^2$ and Δ_z (and as we shall see the probability distributions $\{p^i_z\}_{i\in\mathbb{N}}$) only depend on $\theta - \varphi$.

- we can find the probability distribution p_z^{i+1} of z_{i+1} as the density of the following probability:

$$\mathrm{P}(\{z_{i+1} \in A\}) = \int_0^\infty p_z^i(z) \left(\int_0^\infty \int_0^{2\pi} \mathbb{1}_A(z + \Delta_z(r, \theta, z)) \frac{p_r(r)}{2\pi} dr d\theta \right) dz \tag{7.4.5}$$

  for every $A \subset [0, +\infty)$ Borel measurable. Namely, the probability that the particle's height at the $(i+1)$-th encounter lies within a given set A is obtained by considering the probability the particle was at height z at the i-th encounter, multiplied by the probability that the displacement $\Delta_z(r, \theta, z)$ results in a new height $z + \Delta_z$ falling within A, integrated over all possible values of r, θ and z with their respective probability densities.

We have thus found the probability density p_z^i for every z_i. We recall that, by assumption, the i-th encounter starts exactly at time $(i-1)t_p$, so we can define the probability density $p_z(z, T, z_1)$ for every time $T > 0$ by interpolation. We can also find $\mathrm{MSD}_{xy}(T, z_1)$. Assuming $T = Nt_p$ for some $N > 0$, the particle will have undergone N encounters. Thus:

$$\begin{aligned} \mathrm{MSD}_{xy}(T, z_1) &= \sum_{i=1}^N \int_0^\infty p_z^i(z) \left(n \int_0^\infty \int_0^{2\pi} \|\Delta_{xy}(r, \theta, z)\|^2 r \exp(-nr^2\pi) d\theta dr \right) dz \\ &= \sum_{i=1}^N \int_0^\infty p_z^i(z) \mathbb{E}_{r,\theta}[\|\Delta_{xy}(r, \theta, z)\|^2] dz. \end{aligned} \tag{7.4.6}$$

If T is not an exact multiple of t_p we can define $\mathrm{MSD}_{xy}(T, z_1)$ by interpolation. Namely, if $T \in ((N-1)t_p, Nt_p)$, we define

$$\begin{aligned} \mathrm{MSD}_{xy}(T, z_1) = & \sum_{i=0}^{N-1} \int_0^\infty p_z^i(z) \mathbb{E}_{r,\theta}[\|\Delta_{xy}(r, \theta, z)\|^2] dz \\ & + \frac{|T - Nt_p|}{t_p} \int_0^\infty p_z^{N-1}(z) \mathbb{E}_{r,\theta}[\|\Delta_{xy}(r, \theta, z)\|^2] dz \\ & + \frac{|T - (N-1)t_p|}{t_p} \int_0^\infty p_z^N(z) \mathbb{E}_{r,\theta}[\|\Delta_{xy}(r, \theta, z)\|^2] dz, \end{aligned}$$

where we define $p_z^0 = 0$. Notice that the derivation we followed does not rely on the specific choice of the initial distribution. Indeed, if we replaced a Dirac delta centered at z_1 with a generic distribution p_z^1, the same argument would still apply. We now make the integral in (7.4.6) dimensionless to derive some scaling laws. We choose as non-dimensional variables $\rho = r/\ell$ and $\zeta = z/h$, with h the height of the Stokeslet. As for Δ_{xy}, since the magnitude of the velocity field induced by the Stokeslet is proportional to $|F|/(\ell\mu)$ (with μ the dynamic viscosity of the fluid, see (7.4.2)) we divide and multiply the displacement by $|F|t_p/(\ell\mu) = ft_p/\ell$. We have:

$$\mathrm{MSD}_{xy}(T) = \left[\sum_{i=1}^N \int_0^\infty h p_z^i(h\zeta) \mathrm{I}_{xy}(h\zeta) d\zeta \right] n f^2 t_p^2, \tag{7.4.7}$$

with

$$\mathrm{I}_{xy}(h\zeta) = \left(\int_0^\infty \int_0^{2\pi} \frac{\|\Delta_{xy}(\ell\rho, \theta, h\zeta)\|^2 \ell^2}{f^2 t_p^2} \rho \exp(-n\ell^2\rho^2\pi) d\theta d\rho \right).$$

Since $N = T/t_p$, we can infer that the scaling of $\mathrm{MSD}_{xy}(T)$ behaves like:

$$\mathrm{MSD}_{xy}(T) \approx n f^2 t_p^2 N = n t_p f^2 T.$$

However this estimate can only be valid under certain assumptions, namely that Δ_{xy} is of the order ft_p/ℓ, which is only valid for moderate values of ft_p (see Remark 7.4.2) where the displacement Δ_{xy} can be reasonably estimated as the product of the velocity magnitude and t_p, otherwise there is a non-linear relation between the displacement and the product ft_p. Since in the following we will discuss numerical simulations to validate our model, it is convenient to briefly describe how we implemented (7.4.5). We consider an initial distribution p_z^1. We start by discretising every variable, choosing both a maximum height Z and a maximum distance R. While Z depends on the context (ideally, if we consider a time interval $[0, T]$, it should be high enough such that it is practically impossible for a particle starting in the support of p_z^1 to arrive at height Z in a time T), we can choose R so that the probability of the nearest neighbour being more distant than R is less than a threshold value ε. So, fixing $\varepsilon > 0$ (for instance, we always assume $\varepsilon = 0.001$ in numerical simulations) we have:

$$\mathrm{P}(\{d > R\}) = \exp(-nR^2\pi) \leq \varepsilon \implies R \geq \sqrt{\frac{|\log(\varepsilon)|}{n\pi}} \implies R = \sqrt{\frac{|\log(\varepsilon)|}{n\pi}}.$$

We now explain how we discretise the probability density functions $p_z^i(z)$ and how to find $p_z^{i+1}(z)$ given $p_z^i(z)$ using a finite state Markov chain. First, we partition the interval $[0, Z]$ in K subintervals $I_1, \dots, I_K$ such that

$$[0, Z] = \bigcup_{j=1}^{K} I_j,$$

and we consider the midpoints $\{\bar{z}_j\}_{j=1}^K$ of each interval. We will approximate $\{z_i\}_{i=1}^N$ with a finite Markov chain with states $\{\bar{z}_j\}_{j=1}^K$. We consider a row vector

$$w_1 = \left(w_1^1, \dots, w_1^K\right)$$

of length K, and the mass assigned to each $\bar{z}_j$ is given by

$$w_1^j = \mathbb{P}(\{z_1 \in I_j\}) = \int_{I_j} p_z^1(z)dz.$$

We can now find an approximation of the distribution p_z^2 by working with w_1. We consider the transition probabilities from every state $\bar{z}_i$ to every other state $\bar{z}_j$, obtaining a stochastic matrix Π whose entries are

$$\Pi_{ij} = \text{Probability that, after an encounter, a tracer initially at height } z_i \text{ arrives at height } z_j.$$

We compute the transition matrix in the following way. We consider a mass-splitting map

$$S : [0, Z] \to [0, 1]^k$$

defined as follows:

- If $0 \leq z \leq \bar{z}_1$, then $S(z) = (1, \dots, 0)$;
- if $z \geq \bar{z}_K$, then $S(z) = (0, \dots, 1)$;
- if $z \in [\bar{z}_j, \bar{z}_{j+1}]$ for some $j = 1, \dots, K-1$, then $S(z) = \left(0, \dots, \underbrace{\frac{|S(z) - z_{j+1}|}{|z_{j+1} - z_j|}}_{j\text{-th position}}, \underbrace{\frac{|S(z) - z_j|}{|z_{j+1} - z_j|}}_{(j+1)\text{-th position}}, \dots, 0\right)$.

We can compute each row of the matrix Π as [16]

$$\Pi_i = \int_0^R \int_0^{2\pi} S(z_i + \Delta_z(r, \theta, z_i)) \frac{p_r(r)}{2\pi} d\theta dr. \tag{7.4.8}$$

Given w_i, we can compute w_{i+1} simply as

$$w_{i+1} = w_i \cdot \Pi.$$

[16] We also need to discretise r and φ, but we omit writing it to keep the notation lighter since it is a standard discretisation.

7.4.2 Random encounter times

We now consider a more general model in which the duration of each encounter, previously assumed to be a fixed value t_p, is treated as a random variable with a given probability density p_t. We also assume that the duration of each encounter is identically distributed and independent of all others. While the previous setting could be described in terms of an ensemble of actuators which simultaneously refresh after a given time t_p, it is now more convenient to adopt exclusively the tracer's point of view, assuming that each encounter lasts for a random time t. Once this time has elapsed, the particle immediately falls under the influence of the next nearest actuator, characterised by new encounter parameters (r, θ, z, t). In particular, we do not specify any behaviour of the actuators: we only assume their positions and the direction of the force F they exert to be uniformly distributed.

By the same arguments of the previous case, assuming the tracer's initial position is $(0, 0, z_1)$, we can compute the evolution of the distribution of heights $p_z^i = p_z^i(z, z_1)$, as well as the mean squared displacement in the xy-directions produced during the i-th encounter, which we denote by $\mathrm{MSD}_{xy}^i = \mathrm{MSD}_{xy}^i(z_1)$. As in the previous case, we will omit writing the dependence on z_1 if it is not necessary. Given a probability density p_z^i for the height of the particle at the i-th encounter, we can find the density p_z^{i+1} of the $(i+1)$-th encounter as the density of the probability P defined as:

$$\mathrm{P}(\{z_{i+1} \in A\}) = \int_0^\infty p_z^i(z) \left(\int_0^\infty \int_0^\infty \int_0^{2\pi} \mathbb{1}_A(z + \Delta_z(r, \theta, z, t)) p_\theta(\theta) p_r(r) p_t(t) d\theta dr dt \right) dz \quad (7.4.9)$$

for every $A \subset [0, +\infty)$ Borel measurable. Similarly, we have that

$$\mathrm{MSD}_{xy}^i = \int_0^\infty \int_0^\infty \int_0^\infty \int_0^{2\pi} \|\Delta_{xy}(r, \theta, z, t)\|^2 p_\theta(\theta) p_r(r) p_t(t) p_z^i(z) d\theta dr dt dz. \quad (7.4.10)$$

The mean squared displacement in the xy variables after N encounters, which we will denote by $\mathrm{MSD}_{xy}(N)$, is the sum of the mean squared displacements at each encounter:

$$\mathrm{MSD}_{xy}(N) = \sum_{i=1}^N \mathrm{MSD}_{xy}^i. \quad (7.4.11)$$

Notice that, under our assumptions on the activation times (i.e. that they are independent and identically distributed) mixed term in (7.4.3) vanish by the same argument we have already considered in the case $t = t_p$. Unlike the previous case though, there is now no deterministic relation between the total elapsed time T and the number of encounters N a particle has undergone. However, under our assumptions, we can compute the probability density function of the number of encounters N occurring within a given time T. Indeed, the total time passed after N encounters is given by

$$\tau_N = t_1 + \cdots + t_N,$$

where each t_i is by assumption an independent random variable with distribution p_t. Thanks to the independence assumption we can explicitly compute the density function of τ_N, which is the convolution of N copies of p_t, i.e.:

$$p_{\tau_N} = \underbrace{p_t * \cdots * p_t}_{N \text{ times}}.$$

We can now assign a weight $w_i(T) \in [0, 1]$ to each MSD_{xy}^i according to the probability that a tracer undergoes at least i encounters before time T. We can define such weights as

$$w_i(T) = \mathrm{P}(\{t_1 + \dots t_i < T\}) = \mathrm{P}(\{\tau_i < T\}) = \int_0^T p_{\tau_i}(s) ds. \quad (7.4.12)$$

Thus, we define $\mathrm{MSD}_{xy}(T)$ as

$$\mathrm{MSD_{xy}(T)} = \sum_{i=1}^{\infty} w_i(T)\,\mathrm{MSD}^i_{xy}\,. \tag{7.4.13}$$

Similarly, the probability density function $p_z(z,T)$ of the height of the tracer at time T depends on the number of encounters a particle has undergone. However, the contributions do not simply add up. Therefore, for each N, we estimate the probability that the particle undergoes exactly N encounters by time T, and we give a weight to each p_z^N according to such a probability. Assuming $N > 1$, the probability that the particle has undergone exactly N encounters by time T is:

$$\mathrm{P}(\{t_1 + \dots t_{N-1} < T \text{ and } t_1 + \dots t_N > T\}) = \mathrm{P}(\{\tau_{N-1} < T \text{ and } \tau_{N-1} + t_N > T\}),$$

which can be easily proved to be $w_{N-1}(T) - w_N(T)$. Indeed, since we know that τ_{N-1} has distribution $p_{\tau_{N-1}}$ and t_N has distribution p_t:

$$\begin{aligned}
&\mathrm{P}(\{\tau_{N-1} < T \text{ and } \tau_{N-1} + t_N > T\}) = \mathrm{P}(\{\tau_{N-1} < T \text{ and } t_N > T - \tau_{N-1}\})\\
&= \int_0^T p_{\tau_{N-1}}(t)\left(\int_{T-t}^{\infty} p_t(s)ds\right)dt = \int_0^T p_{\tau_{N-1}}(t)\left(1 - \int_0^{T-t} p_t(s)ds\right)dt\\
&= \underbrace{\int_0^T p_{\tau_{N-1}}(t)dt}_{=w_{N-1}(T)} - \int_0^T p_{\tau_{N-1}}(t)\int_0^{T-t} p_t(s)dsdt\\
&= w_{N-1}(T) - \int_0^T p_{\tau_{N-1}}(t)\int_t^T p_t(s-t)dsdt\\
&= w_{N-1}(T) - \int_0^T\int_0^T p_{\tau_{N-1}}(t)p_t(s-t)dtds\\
&= w_{N-1}(T) - \int_0^T \underbrace{\left(\int_{\mathbb{R}} p_{\tau_{N-1}}(t)p_t(s-t)dt\right)}_{=p_{\tau_{N-1}}*p_t(s)=p_{\tau_N}(s)} ds = w_{N-1}(T) - w_N(T).
\end{aligned}$$

If $N = 1$, we have to account for the possibility that the first encounter lasts more than T. That is:

$$\mathrm{P}(\{t_1 \geq T\}) = 1 - \mathrm{P}(\{t_1 < T\}) = 1 - w_1(T) = w_0(T) - w_1(T),$$

where we set by definition $w_0(T) = 1$. Thus, we define:

$$p_z(z,T) := \sum_{i=1}^{\infty}(w_{i-1}(T) - w_i(T))p_z^i(z). \tag{7.4.14}$$

We also note that the same non-dimensionalisation as in (7.4.7) can be performed, substituting t_p with the expected duration of an encounter $\mathbb{E}[t]$ and estimating the number of encounters in a time T to be $\approx T/\mathbb{E}[t]$.

7.4.3 Evolution of a density advected by randomly forced velocity fields

In this subsection we present some ways to pass from the Lagrangian point of view to a Eulerian one, approximating the average evolution of a density (or, equivalently, of the probability density of the position of a tracer) evolving under the stirring mechanism we considered. We present three possible approaches, assuming the density is initially concentrated at a point mass at $\vec{r}_1 = (x_1, y_1, z_1)$. The density at time t will be denoted by $u(t, \vec{r}, \vec{r}_1)$, with $\vec{r} = (x, y, z)$. If the initial condition is not a Dirac delta centered at $\vec{r}_1$, but rather a diffuse density u_0 in $\mathbb{R}^2 \times [0, +\infty)$, we can find the density $\bar{u}(t, \vec{r})$ at time t simply by superposition, i.e.:

$$\bar{u}(t,\vec{r}) = \int_{[0,+\infty)\times\mathbb{R}^2} u_0(\vec{r}_1)u(t,\vec{r},\vec{r}_1)d\vec{r}_1. \tag{7.4.15}$$

We also notice that, by isotropy in the xy variables, it holds that

$$u(t, \vec{r}, \vec{r}_1) = u(t, (x - x_1, y - y_1, z), (0, 0, z_1)), \tag{7.4.16}$$

i.e. we can assume without loss of generality that $\vec{r}_1 = (0, 0, z_1)$. It would be interesting to derive a density u using a purely Eulerian approach. In this case, u might be obtained as the solution of a PDE of the form

$$\begin{cases} \partial_t u = F(z, Du, D^2 u) \\ u(0, \cdot) = u_0 \end{cases}. \tag{7.4.17}$$

At the moment, we do not know how to derive a density u from a PDE like (7.4.17), but this is still a work in progress.

First approach: Markov chain

The first approach is based on extending the argument previously applied to the variable z to the pair (ρ, z), where $\rho = \sqrt{x^2 + y^2}$ denotes the distance from the origin in the xy-plane. Since the tracer particle is initially located at $\vec{r}_1 = (0, 0, z_1)$, isotropy in the xy variables implies that the density $u(t, \vec{r}, \vec{r}_1)$ is radially symmetric and can therefore be represented as a function of $(\rho, z) = \left(\sqrt{x^2 + y^2}, z\right)$. We then treat (ρ, z) as a Markov chain with initial state $(0, z_1)$. Given the distribution $p^i_{(\rho,z)}$ at the i-th encounter, the distribution at the $(i+1)$-th encounter is obtained as the density of the following probability:

$$\begin{aligned} &\mathrm{P}(\{(\rho_{i+1}, z_{i+1}) \in A \times B\}) \\ &= \int_{[0,+\infty)^4} \int_0^{2\pi} \mathbb{1}_A(\|(\rho, 0) + \Delta_{xy}(r, \theta, z, t)\|) \mathbb{1}_B(z + \Delta_z(r, \theta, z, t)) p_\theta(\theta) p_r(r) p_t(t) p^i_{(\rho,z)}(\rho, z) d\theta dr dt dz d\rho. \end{aligned}$$

Notice that, by isotropy, at each step we can assume that if the tracer is at distance ρ from the origin, its xy coordinates are $(\rho, 0)$. The density at a given time t can be found arguing as for (7.4.14), i.e.

$$p_{(\rho,z)}(t, \vec{r}, \vec{r}_1) = \sum_{i=1}^{\infty} [w_{i-1}(t) - w_i(t)] p^i_{(\rho,z)}(\rho, z, \vec{r}_1), \tag{7.4.18}$$

with the weights $w_i(t)$ defined as in (7.4.12). By radial symmetry we define $u(t, \vec{r}, \vec{r}_1)$ as

$$u(t, \vec{r}, \vec{r}_1) = \frac{p_{(\rho,z)}\left(t, \sqrt{x^2 + y^2}, z, z_1\right)}{2\pi\sqrt{x^2 + y^2}}. \tag{7.4.19}$$

In general, if $\vec{r}_1 = (x_1, y_1, z_1)$, the density u can be found by (7.4.16), i.e.

$$u(t, \vec{r}, \vec{r}_1) = \frac{p_{(\rho,z)}\left(t, \sqrt{(x - x_1)^2 + (y - y_1)^2}, z, (0, 0, z_1)\right)}{2\pi\sqrt{(x - x_1)^2 + (y - y_1)^2}}. \tag{7.4.20}$$

This approach might be computationally demanding, since it is necessary to compute the transition probabilities from every possible state (ρ, z) to every other possible state (ρ', z'). In particular the transition matrix Π may be sparse, but the density u has to be supported within the domain $I_{(x_1,y_1)} \times [0, Z]$, with $I_{(x_1,y_1)}$ a neighbourhood of (x_1, y_1).

Second approach: Monte Carlo

The second approach relies on a Monte Carlo approximation of the probability density function of (ρ, z). After discretising the encounter parameters (r, θ, z, t), we save the displacements $\Delta(r, \theta, z, t)$ with respect to every possible set of encounter parameters. Then, we proceed as follows. Given the initial position $\vec{r}_1 = (0, 0, z_1)$, we select the values of r, θ, t according to their respective probability densities. The tracer will then be displaced to $(x_2, y_2, z_2) = \vec{r}_1 + \Delta(r, \theta, z_1, t)$. We compute the value of $\rho_2 = \sqrt{x_2^2 + y_2^2}$ and iterate the procedure up to the N-th encounter. Averaging over many realisations gives an approximation of the probability density $p^i_{(\rho,t)}$ of the couple (ρ_i, z_i) at the i-th encounter for every $i = 1, \dots, N$, which arguing as in (7.4.18) gives the density $p_{(\rho,z)}(t, \vec{r}, \vec{r}_1)$ and thus an approximate density $u(t, \vec{r}, \vec{r}_1)$ defined as in (7.4.19) and (7.4.20). We remark that, a priori, a Monte Carlo approach can be used also to approximate $\mathrm{MSD}_{xy}(t, z_1)$ and $p_z(t, z, z_1)$, that we explicitly derived in the previous subsection.

Third approach: martingale central limit theorem

We consider a tracer particle initially placed at $\vec{r}_1 = (0, 0, z_1)$. The third approach does not require a direct approximation of the probability density function of $\rho = \sqrt{x^2 + y^2}$, since it only makes use of the mean squared displacement in the xy variables after N encounters, i.e. $\mathrm{MSD}_{xy}(N, z_1)$, defined in (7.4.11). Its applicability, however, is limited to the long-time regime, where many encounters between the tracer and an actuator have occurred. Moreover, we approximate only the probability density $u(t, (x, y), \vec{r}_1)$ of the tracer's position projected onto the xy-plane, thus neglecting the variable z. This approach is based on an extension of the classical central limit theorem for independent and identically distributed random variables to the case of martingales.

Definition 7.4.1. ***Martingale, [15]***
A $\mathbb{R}^d$-valued process $M = (\Omega, \mathcal{F}, (\mathcal{F}_i)_{i\in\mathbb{N}}, (M_i)_{i\in\mathbb{N}}, \mathrm{P})$ is a (discrete time) martingale if for every $i, j \in \mathbb{N}$ such that $i \geq j$ it holds

$$\mathbb{E}[X_i \mid \mathcal{F}_j] = X_j, \tag{7.4.21}$$

where $\mathbb{E}[X_i \mid \mathcal{F}_j]$ denotes the conditional expectation *with respect to $\mathcal{F}_j$.*

We state here a non-sharp version of the theorem, which suffices for our purposes. A more general formulation can be found in [90, Theorem 3.2].

Theorem 7.4.1. ***Martingale central limit theorem, [90, Theorem 3.2]***
Let $\{X_n\}_{n=0}^\infty$ be a $\mathbb{R}^d$-valued martingale with bounded increments and such that $X_0 = 0$. Let

$$\sigma_n^2 := \mathbb{E}[\|X_n - X_{n-1}\|^2 \mid X_{n-1}, \dots, X_0] \quad \text{with } \sigma_0^2 := 0,$$

and assume that the components X_n^i are independent and identically distributed, with $\mathbb{E}[X_n^i] = 0$, for every $i = 1, \dots, d$ and for every $n \geq 0$. Let

$$v_n := \frac{1}{d} \sum_{i=0}^{n} \sigma_i^2.$$

Then, $X_n / \sqrt{v_n}$ converges in distribution to the d-dimensional Gaussian distribution $N(0, I)$ as $n \to +\infty$.

The total displacement S_N in the xy variables of a tracer particle after N encounters is a martingale, which satisfies the assumptions of Theorem 7.4.1. Indeed, considering the natural σ-algebra $\mathcal{F}_j = \sigma(S_j, \dots, S_1)$ generated by the process, condition (7.4.21) is equivalent to

$$\mathbb{E}[S_N - S_{N-1} \mid \mathcal{F}_{N-1}] = 0 \quad \forall N \geq 1.$$

Since $S_N = \sum_{i=1}^N \Delta^i_{xy}$, we have that $E[S_N - S_{N-1} \mid \mathcal{F}_{N-1}] = E[\Delta^N_{xy} \mid \mathcal{F}_{N-1}] = 0$, because knowing all the previous history of S_N does not yield any information on the N-th step [17], whose direction is uniformly

[17] To be precise, we need to use the *tower property* of the conditional expectation. Let R_N be the norm of the N-th displacement and let U_N be its direction. Since the norm of the displacement depends on the height of the particle, R_N is not independent of $\mathcal{F}_{N-1}$. However:

$$\begin{aligned}\mathbb{E}[\Delta^N_{xy} \mid \mathcal{F}_{N-1}] &= \mathbb{E}[R_N U_N \mid \mathcal{F}_{N-1}] = \mathbb{E}[\mathbb{E}[R_N U_N \mid \sigma(\mathcal{F}_{N-1}, R_N)], \mathcal{F}_{N-1}] \\ &= \mathbb{E}[R_N \mathbb{E}[U_N \mid \sigma(\mathcal{F}_{N-1}, R_N)], \mathcal{F}_{N-1}] = \mathbb{E}[R_N \underbrace{\mathbb{E}[U_N]}_{=0}, \mathcal{F}_{N-1}] = 0.\end{aligned}$$

distributed. We assume that the time T is sufficiently large, so that the minimum number of encounters is high enough to apply Theorem 7.4.1, where

$$v_N = \frac{\mathrm{MSD}_{xy}(N, z_1)}{2}.$$

We therefore approximate the probability distribution of the position (in the xy variables) of the tracer after N encounters, assuming $\vec{r}_1 = (0, 0, z_1)$, as

$$\Phi^N(x, y, z_1) = \frac{1}{\pi\, \mathrm{MSD}_{xy}(N, z_1)} \exp\left(-\frac{x^2 + y^2}{\mathrm{MSD}_{xy}(N, z_1)}\right),$$

with $\mathrm{MSD}_{xy}(N, z_1)$ defined as in (7.4.11). From the distributions Φ^N, we can argue as in (7.4.14) and (7.4.18) to approximate the distribution at time t as

$$\Phi(t, x, y, z_1) = \sum_{i=1}^{\infty}[w_{i-1}(t) - w_i(t)]\Phi^N(x, y, z_1), \quad \text{with } w_i(t) \text{ defined as in (7.4.12).}$$

We approximate the probability distribution at time t with

$$u(t, (x, y), \vec{r}_1) = \Phi(t, x, y, z_1).$$

In the general case of $\vec{r}_1 = (x_1, y_1, z_1)$, by (7.4.16) we have:

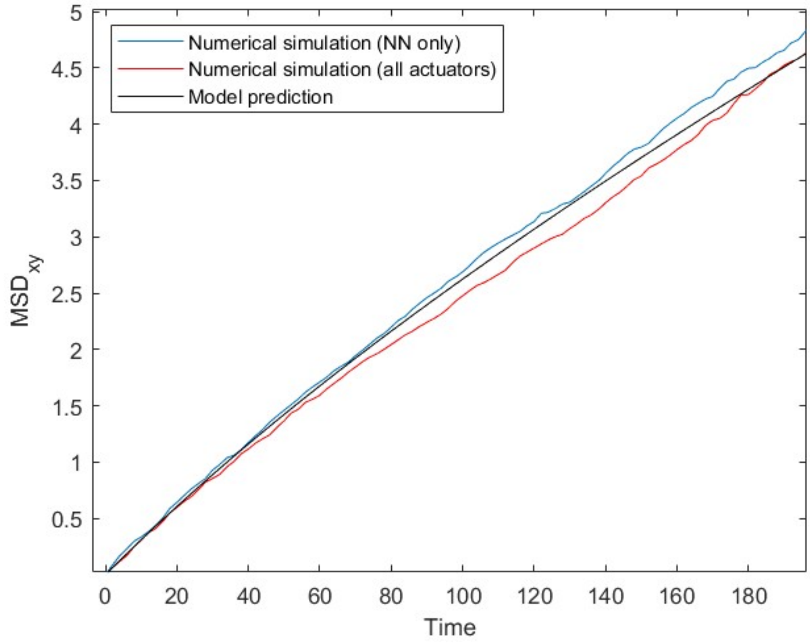
$$u(t, (x, y), \vec{r}_1) = \Phi(t, x - x_1, y - y_1, z_1).$$

7.4.4 Numerical experiments

In this final subsection, we present the results of numerical simulations that support the validity of the model. We start with the description of the numerical simulations we used to validate the case with a fixed encounter time t_p. We will leave for later the description of the benchmark simulations for the more general case of random encounter times.

At time $t = 0$, we generate N_p actuators whose positions are uniformly distributed within a disc of radius R centered at the origin, and whose forcing directions are also uniformly distributed. The radius R is chosen so that $R^2\pi n = N_p$, for a prescribed number density n. Each actuator is modeled as a regularised Stokeslet with radius δ, located at a height h above a no-slip floor, which exerts a force F parallel to the ground on the fluid. The velocity field generated by each actuator is defined according to the formulas derived in [2], with all Stokeslets having the same strength f (see Remark 7.4.1). After a time interval t_p, which we will refer to as *activation time*, the entire carpet of actuators is regenerated following the same rules described above. We also specify the dynamics followed by a tracer, whose initial position is $(0, 0, z_1)$ for a given initial height $z_1 > 0$. The tracer is either advected by the velocity field resulting from the collective action of all N_p actuators, or it moves under the flow generated by its nearest neighbour only (determined at the time of carpet formation and assumed not to change during the particle's motion). We will show that, under the assumption of small number density n, the two dynamics are essentially the same.

We will keep some of the quantities fixed in almost all numerical simulations, if not otherwise stated. In particular, we will always implicitly assume $\delta = 0.215, h = 0.25$ and $f = 30$. These values are not arbitrary, as we will see they derive from real measurements on a population of *Stylonychia* [155]. We can see in Figure 7.11 that the mean squared displacement and the height distribution produced by the collective action of the actuators and by the nearest neighbour only are practically the same, and they are in good agreement with the model predictions.



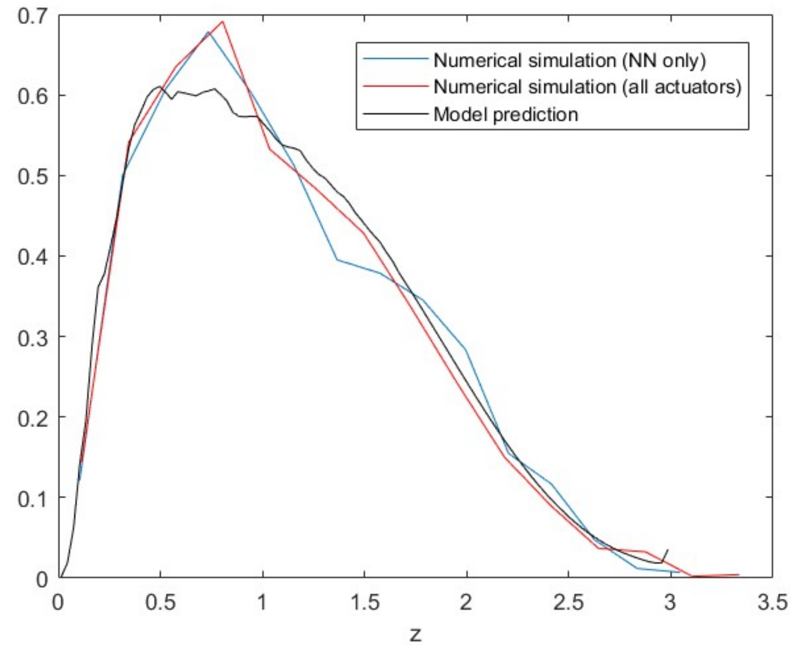
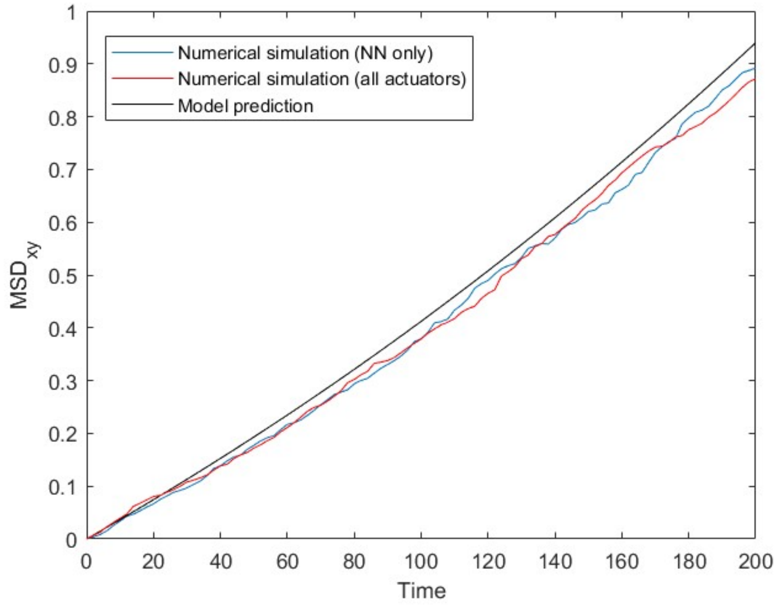


Figure 7.11: (Up) The mean squared displacement in the xy variables (MSD_{xy}) up to time $T = 200\,\mathrm{s}$ considering a number density $n = 0.005$, a fixed activation time $t_p = 2\,\mathrm{s}$ and an initial height for the tracer $z_1 = 0.75$. We compare the model prediction (in black) with the MSD_{xy} obtained by averaging 2000 realisations where the tracer particle is either advected by its nearest actuator only (in blue) or by the collective flow (in red) produced by $N_p = 200$ actuators which refresh simultaneously every t_p seconds. (Down) The distribution of heights after $T = 200\,\mathrm{s}$ predicted by the model (in black) or by averaging over 2000 realisations of the numerical simulations just described.

We can also examine how the initial height of the tracer affects the mean squared displacement. In Figure 7.12 we see that, for a starting height of $z_1 = 0.05$, the mean squared displacement is four times smaller than for a particle starting at $z_1 = 0.75$. We also observe that most of the times a tracer particle remains near the floor, with probability densities very concentrated near $z_1 = 0.05$. This is caused by the no-slip boundary condition. Once again, we see how the model predicts with good accuracy both the mean squared displacement and the distribution of the heights.



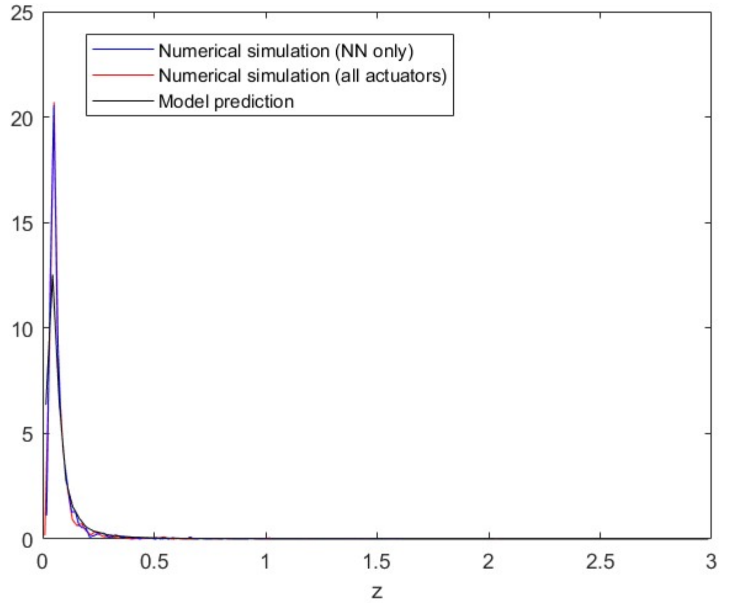
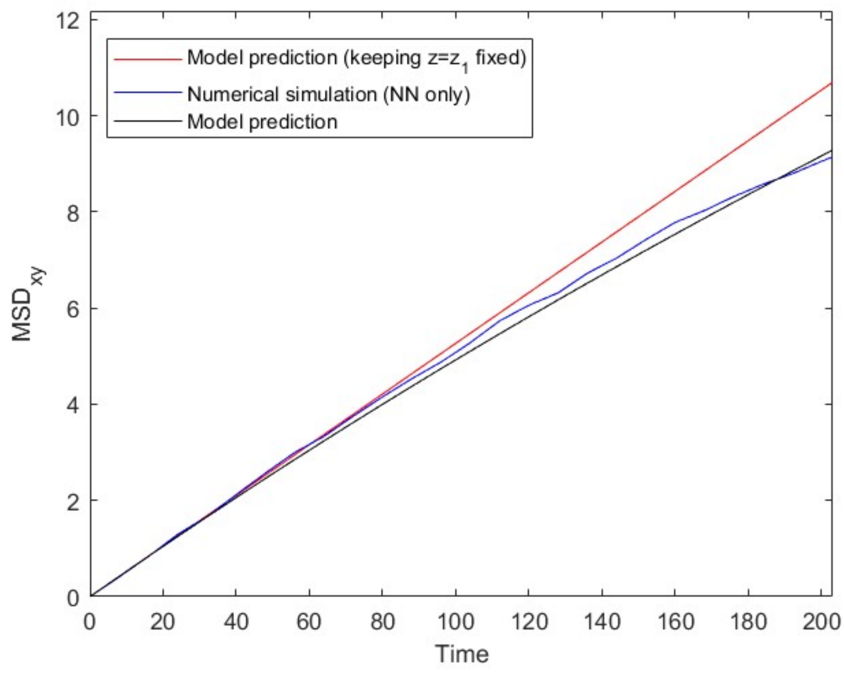


Figure 7.12: (Up) The mean squared displacement in the xy variables (MSD_{xy}) up to time $T = 200\,\mathrm{s}$ considering a number density $n = 0.005$, a fixed activation time $t_p = 2s$ and an initial height for the tracer $z_1 = 0.05$. We compare the model prediction (in black) with the MSD_{xy} obtained by averaging 2000 realisations where the tracer particle is either advected by its nearest neighbour only (in blue) or by the collective flow (in red) produced by $N_p = 200$ actuators which refresh simultaneously every t_p seconds. (Down) The distribution of heights after $T = 200\,\mathrm{s}$ predicted by the model (in black) or by averaging over 2000 realisations of the numerical simulations just described. The blue and the red lines, in particular, are almost indistinguishable.

One might wonder if the displacements in the z variable can be neglected and if we could estimate the mean squared displacement after a time T simply assuming $z = z_1$ fixed. In Figure 7.13, we can see that in general this is not the case. When T is moderate, assuming a fixed height might provide a good estimate of $\mathrm{MSD}_{xy}(T)$; however, for larger T, this is not true. The problem is that, after a sufficiently long time T, the particle will, on average, have spread over different heights. In particular, when particles are far from the floor the displacement they experience becomes smaller. As time passes, the slope of the curve changes, so MSD_{xy} and T are, in general, not linearly dependent as in (7.2.4).



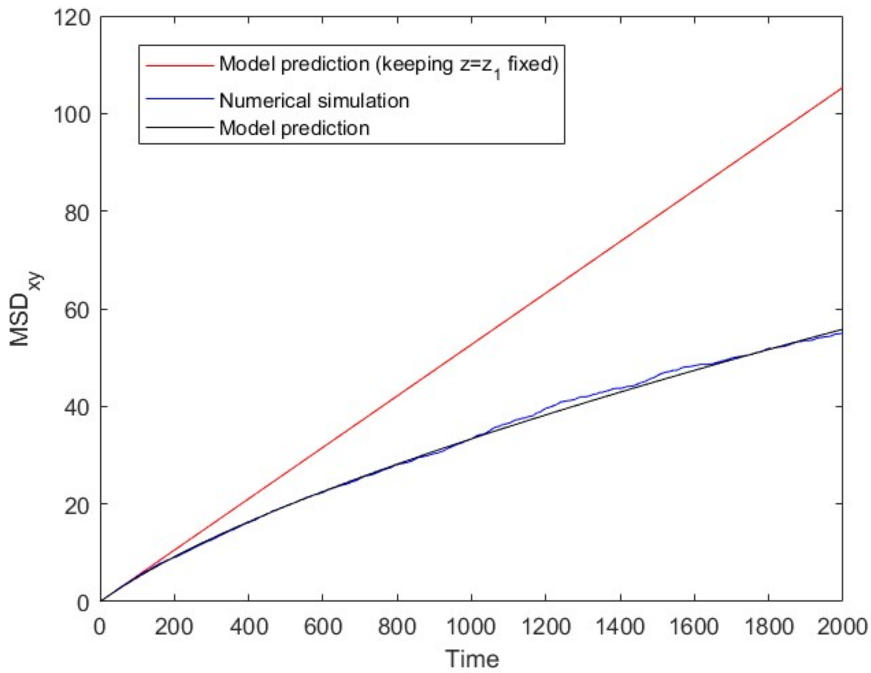


Figure 7.13: (Up) The mean squared displacement in the xy variables (MSD_{xy}) up to time $T = 200\,\mathrm{s}$ considering a number density $n = 0.005$, a fixed activation time $t_p = 8s$ and an initial height $z_1 = 0.75$. We compare the model prediction (in black) with the prediction by the same model assuming the height $z = z_1$ to be constant (in red). We also plot the MSD_{xy} obtained by averaging over 2000 realisations where the tracer particle is advected by its nearest neighbour only (in blue), considering $N_p = 200$ actuators which refresh simultaneously every t_p seconds. (Down) The same as the upper graph, with the only difference that $T = 2000\,\mathrm{s}$. We can see how assuming $z = z_1$ does not give a good prediction of MSD_{xy} for large times.

The model manages to predict with good precision the height distribution even after long times, see Figure 7.14.

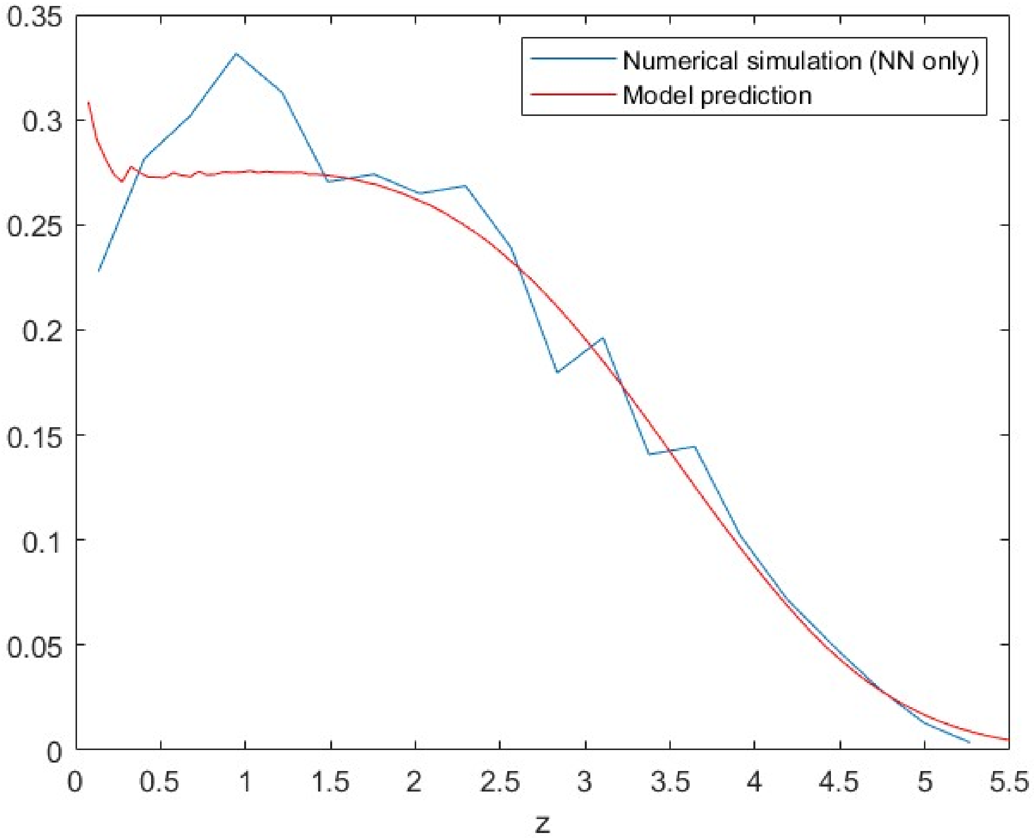


Figure 7.14: Height distributions of a tracer after $T = 2000\,\mathrm{s}$, assuming a number density $n = 0.005$, a fixed activation time $t_p = 2\,\mathrm{s}$ and an initial height $z_1 = 0.75$. We compare the model prediction (in red) with the distribution obtained averaging over 2000 realisations where the tracer particle is advected by its nearest neighbour only (in blue), considering $N_p = 200$ actuators which refresh simultaneously every t_p seconds.

The effect of the initial height can also be seen in Figure 7.15, where we compute equivalent diffusion coefficients after various periods of time and assuming different initial heights of the tracer. By "equivalent diffusion coefficient" we mean a diffusion coefficient ν_T that yields, after a given time T, the same mean squared displacement. We can clearly see from Figure 7.15 that equivalent diffusion coefficients depend more on starting heights for moderate values of T, but for longer times this difference becomes smaller and smaller. Since in our model the height z is treated as a Markov chain, its probability density converges, as time passes, to its (unique) stationary probability distribution [18]. Thus, it is not surprising that diffusion coefficients tend to converge to the same value as time passes, since the height distribution will converge to the stationary one. In any case, we recall that our model is based on the assumption that the effect of the nearest neighbour on the tracer is dominant, and this can only be valid assuming the tracer is near the floor. Otherwise, a far-field approach where the action of all the actuators is considered is more suited [89].

[18]This is a classical result: irreducible and aperiodic Markov chains are *ergodic*, and they have the property of having a unique stationary distribution to which they converge for every initial datum.

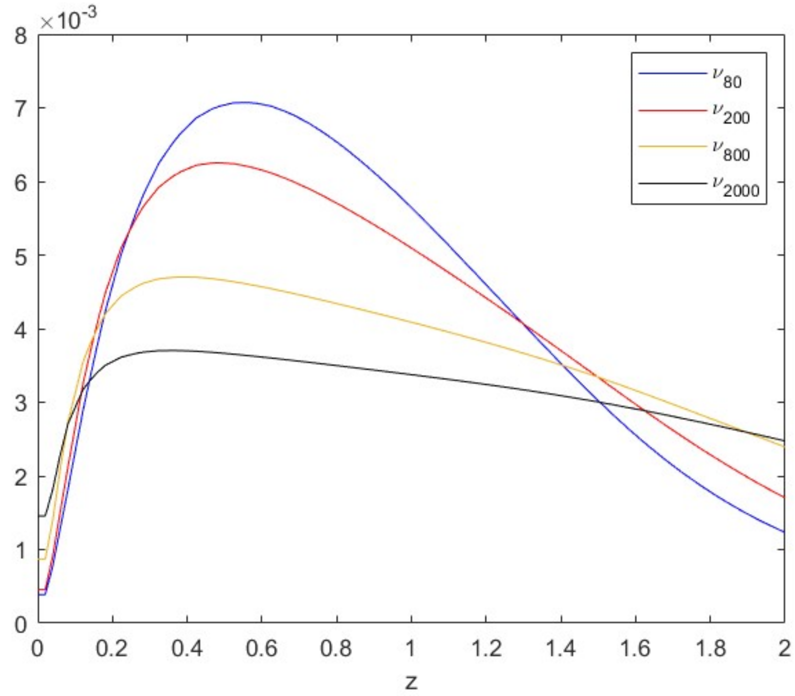


Figure 7.15: Different equivalent diffusion coefficients, assuming $n = 0.005$ and $t_p = 2\,\mathrm{s}$, computed using the relation $\mathrm{MSD}_{xy}(T) = 4\nu_T T$, where we use our model to compute $\mathrm{MSD}_{xy}(T)$. We consider different initial heights z, and for each starting height we compute the equivalent diffusion coefficient ν_T after a time T. We consider $T = 80\,\mathrm{s}$ (in blue), $T = 200\,\mathrm{s}$ (in red), $T = 800\,\mathrm{s}$ (in yellow) and $T = 2000\,\mathrm{s}$ (in purple). We can see how, as T becomes larger, the effect of the starting height on the equivalent diffusivity gradually decreases. We discretise the height in an interval $[0, 10]$ to avoid possible interferences of the upper boundary.

We now employ our model to estimate the enhanced diffusion generated by a feeding population of *Stylonychia*. In contrast to the previous numerical experiments, we can rely on real laboratory data as a benchmark. Before proceeding, we briefly summarise the experimental setup described in [155], which provides the benchmark data we use.

The experiment was performed on a culture of *Stylonychia* within a Petri dish whose glass plates were separated by a $H = 3\,\mathrm{mm}$ gap. As we will see, *Stylonychia* typically exert forces at a height $h \approx 50\,\mathrm{\mu m}$, and for values $H/h > 8$ the effect of the ceiling on the flow can be neglected [119]. The movements and the direction of the force F of each *Stylonychia* are measured using specialised laboratory equipment and MATLAB programs. See [155, Supporting Information, Section 2.A] for more details.

We consider the Dataset S01 provided in [155, Supporting Information]. Such data contain the xy coordinates and the direction of the force F of 485 *Stylonychia*, sampled at time intervals $\Delta t \approx 0.043\,\mathrm{s}$. The data are quite noisy: coordinates are often saved as NaN (i.e. not a number). The reason is that, at the moment the coordinates were measured, the microorganism was moving too fast, exceeding a threshold velocity of $v_{\max} = 500\,\mathrm{\mu m\,s^{-1}}$. This value was established in [155] as the threshold speed above which microorganisms act on the fluid as moving stresslets rather than Stokeslets, whose effects are negligible compared to still actuators which act with a non-zero net force. For this reason, we decide to only retain coordinates of *Stylonychia* which do not "move too much", in particular we will only consider *Stylonychia* whose coordinates have less than 30% of NaN. From now on, we will call *actuators* the 176 members of this specific subset of the population of microorganisms, neglecting the others. For each actuator, we consider its xy coordinates, and when a NaN value is encountered we assign a position using linear interpolation between the nearest non-NaN value before and the nearest non-NaN value after it. We perform the same interpolation for the directions of the force. We thus have a set of xy coordinates and force directions for a subset of the *Stylonychia*, which we model as regularised Stokeslets. The benchmark for testing the model consists of computing the mean squared displacement in the xy variables and the distribution of the heights at the final time $T = 210\,\mathrm{s}$ of a tracer particle advected by the collective action of the actuators. Before running the numerical simulations, there are some issues we need to address. First, the number density of the actuators we consider may be too high for our model. We always crucially assumed a low number density, so that interactions between actuators could be neglected. As we can see in Figure 7.16, where we consider the initial positions of the actuators, they seem too crowded. Moreover, we cannot always consider a tracer with a fixed initial position $(0, 0, z_1)$, since positions and orientations of *Stylonychia* are given and the displacement would be deterministic. We solve both of these problems as follows. In the numerical simulations, we will only consider actuators which lie inside a given circle of radius R and center C for the whole time $T = 210\,\mathrm{s}$. For each simulation, we randomly select M actuators chosen between the ones which stay inside the circle, and we only consider the flow they generate. Moreover, to increase randomness, we consider the tracer's initial position to be (x, y, z_1), with z_1 fixed and (x, y) uniformly chosen in a circle with center C and radius $r < R$.

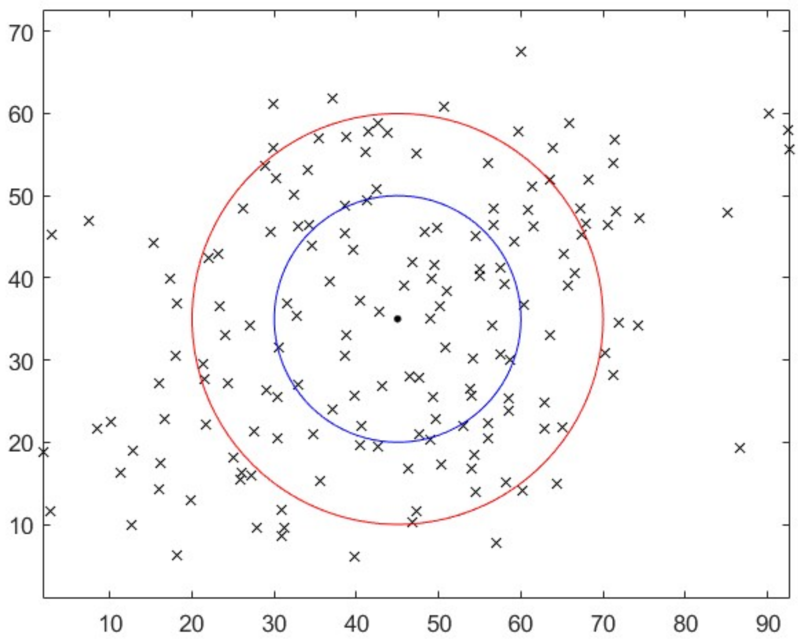


Figure 7.16: The black crosses denote the initial positions of the actuators. In numerical simulations, we select randomly $M = 10$ of the 46 actuators which always stay inside the red circle, with center $C = (45, 35)$ and radius $R = 25$. We also choose a random starting position (with a prescribed height) for the tracer particle inside the blue circle, which has center C and radius $r = 15$. With these choices, the number density of the actuators in numerical simulations is $n \approx 0.005$.

We now need to specify the parameters of the actuators, namely the strength f of the regularised Stokeslets they generate, its radius δ and the height h at which thy exert the force F. Such parameters are measured in [155] and they are (we will neglect the measure uncertainty and directly consider the mean value):

$$F = 1200\,\mathrm{pN}, \quad h = 50\,\mu\mathrm{m}, \quad \delta = 43\,\mu\mathrm{m}.$$

We will use, as unit of measure for length, the typical body size of the actuator, which is $L_c = 200\,\mu\mathrm{m}$. Since we are interested in the strength f of the Stokeslet, we also need to divide F by the dynamic viscosity of water, i.e. $10^{-3}\,\mathrm{Pa\,s}$. We find the values (we omit writing the units of measurement):

$$f = 30, \quad h = 0.25, \quad \delta = 0.215.$$

We are now ready to run simulations, but we still miss some of the data needed for our model. Namely, we need to specify the probability distributions p_r, p_θ and p_t. As for p_r, since we randomly select a fixed number M of actuators in a circle of radius R, we can assume p_r to be the probability distribution of the distance from the nearest neighbour (7.4.4) assuming a number density $n = M/(R^2\pi)$, while for p_θ, since we assume isotropy, we consider a uniform distribution. In order to find p_t, we first need to specify when we consider an actuator to stop being active, i.e. to stop exerting a force in a given direction. There are two events which we consider to stop this process:

1. The actuator moves too fast, i.e. it is swimming away from its current position;
2. the actuator changes the direction of the force. Indeed microorganisms as *Stylonychia*, when surrounded by food, perform so-called *side stepping reactions* (SSR). These are fast movements where the microorganism does not change position, but reorients itself in order to pump water into its feeding apparatus from a different direction.

In our model, we consider both of these events to interrupt an encounter, i.e. the motion of a tracer particle. We also assume, for simplicity, that activation times can only be natural numbers. Then we simply look at positions and orientations $1\,\mathrm{s}$ apart. If the actuator has not moved more than 2.5 (we recall that the threshold velocity is $v_{\max} = 500\,\mu\mathrm{m\,s^{-1}}$ and the unit of length we consider is the typical size of a *Stylonychia*, i.e. $L_c = 200\,\mu\mathrm{m}$) and the angle of the force dos not change "too much" (we consider a threshold value of $\pi/6$, since SSR often change the angle more than that, see [155]), we assume it is still active. When one of the two events occur, we reset the time. In this way, we can find an empirical distribution $p_T(T)$ for the total activation time T of an actuator. We point out that this distribution is *not* the one needed in our model. Indeed, we need to find the distribution p_t of the duration t of an encounter from the point of view of the tracer, while p_T is the distribution of the total activation time from the point of view of an actuator. The distribution p_t can however be derived from p_T. First, we derive the conditional probability $\hat{p}_T$ of total activation times for actuators which we assume to have an encounter with the tracer. To summarise:

- p_t is the distribution of encounter durations from the point of view of the tracer;
- p_T is the distribution of the *total* time an actuator stays active, regardless of the tracer;
- $\hat{p}_T$ is the distribution of the *total* time an actuator stays active, conditioned on having an encounter with the tracer.

The distribution $\hat{p}_T$ is *length-biased* [134], i.e. a tracer is more likely to encounter actuators which stay active for longer, and it is proportional both to T and to $p_T(T)$. Normalising the probability, we get

$$\hat{p}_T(T) = \frac{T p_T(T)}{\mathbb{E}[T]}\text{ , with } \mathbb{E}[T] = \int_0^\infty T p_T(T) dT.$$

We can now derive p_t from its cumulative distribution. Namely, let s be the duration of an encounter and let t be the total time that the actuator associated to such encounter stays active. Then, given $S, T > 0$:

$$\mathrm{P}(\{s > S\}) = \int_S^\infty \mathrm{P}(\{s > S | t = T\})\, \mathrm{P}(\{t = T\}) dT = \int_S^\infty \frac{T - S}{T} \hat{p}_T(T) dT.$$

Differentiating with respect to S the cumulative function $\mathrm{P}(\{s \leq S\}) = 1 - \mathrm{P}(\{s > S\})$, we find that

$$p_t(t) = \frac{\int_t^\infty p_T(T) dT}{\mathbb{E}[T]}.$$

We can see in Figure 7.17 and Figure 7.18 a comparison between the model prediction and numerical simulations, which are in good agreement, assuming different starting heights $z_1 = 0.25$ and $z_1 = 0.75$ for the tracer. In the figures we also report linear fits to represent the slope of the mean squared displacement MSD_{xy} predicted by the model (constants are added to avoid overlaps). According to (7.2.4) the slopes, divided by 4, give the enhanced diffusion coefficient κ generated by the action of the actuators. Converting the results to standard units of measurement (we recall that we assumed the characteristic size L_c of a *Stylonychia* as the unit of length, with $L_c = 200\,\mu\mathrm{m}$), we obtain

$\kappa \approx 3.8 \times 10^{-10}\,\mathrm{m^2\,s^{-1}}$ if $z_1 = 0.25 L_c = 50\,\mu\mathrm{m}$ and $\kappa \approx 6 \times 10^{-10}\,\mathrm{m^2\,s^{-1}}$ if $z_1 = 0.75 L_c = 150\,\mu\mathrm{m}$.

It is numerically computed [155, Section 5.A] that the ensemble of *Stylonychia* our dataset refers to, assuming tracers initially placed at height $z_1 = 150\,\mu\mathrm{m}$, generates an enhanced diffusion whose maximum diffusion coefficient is $D_{Flxy} \approx 6 \times 10^{-9}\,\mathrm{m}^2\,\mathrm{s}^{-1}$. This result seems in accordance with our predictions: the enhanced diffusion we compute assuming $z_1 = 150\,\mu\mathrm{m}$ is ten times smaller than D_{Flxy}, but this was to be expected since in each realisation we only consider a small subset of actuators. We also test our model assuming a higher number density $n = 0.02$, randomly selecting 40 actuators instead of 10 for each realisation, see Figure 7.19. Since the slopes of the mean squared displacement curves obtained from numerical experiments and from the model differ, the enhanced diffusion predicted by the model is incorrect. This discrepancy suggests that the number density is likely too high for the model's assumptions to remain valid. Finally, in Figure 7.20 we consider the full experiment analysed in [155] [19], without selecting additional subsets of actuators. However, we remark that we shall still select only actuators which do not "move too much". As we already noticed, this results in the total number of *Stylonychia* considered to drop from 485 to 176. To compute the number density n of actuators to assume in our model when considering the action of all 176 actuators, we exploit the fact that, under the assumption of a uniform distribution of microorganisms, the number density n and the average minimum distance $\mathbb{E}[r]$ between actuators satisfy the following relation:

$$n = \frac{1}{4\mathbb{E}[r]^2}. \tag{7.4.22}$$

We compute that the average minimum distance between the 176 actuators considered is $\mathbb{E}[r] \approx 2.9$, which gives $n \approx 0.03$. Using such a value in our model, we find $\mathrm{MSD}_{xy}(t)$ as in Figure 7.20. We also plot linear fits approximating the slope of $\mathrm{MSD}_{xy}(t)$ at different times, since its behaviour is not linear. The resulting diffusion coefficient κ is in the range $1.2 \times 10^{-9} - 2.6 \times 10^{-9}\,\mathrm{m}^2\,\mathrm{s}^{-1}$. While smaller than the coefficient D_{Flxy} computed in [155], we believe that such a mismatch is justifiable. Indeed:

- We only considered a fraction of actuators, neglecting those that were "moving too much", where we imposed an arbitrary limit to how much we allowed an actuator to move before neglecting it (i.e. we did not considered actuators whose position vectors contained more than 30% NaN entries). Although we justified this assumption by noting that, above a threshold speed $v_{\max}$, the effect of a crawling *Stylonychia* can be approximated as that of a moving stresslet (whose influence on the fluid is negligible compared to static Stokeslets), we did not consider the possibility that the microorganism's speed might greatly exceed $v_{\max}$. In such a case, the drag force could exceed the force exerted on the fluid by the feeding *Stylonychia* [155, Figure 3D], so that the overall effect is that of a moving Stokeslet, which might not be negligible;

- a crucial assumption of our model is that a tracer particle is effectively displaced by one actuator at a time, with two actuators never interacting with each other. While this assumption is justifiable assuming a small number density, it is not when the actuators are too crowded. The situation considered in [155] appears to correspond to the latter case, since hydrodynamic interactions between individual feeding flows is observed [155, Section 4] and Figure 7.19 shows that even for $n = 0.02$ the model fails to accurately predict the mean squared displacement.

To conclude, we pose two open questions.

1. What is the critical height Z_c above which our model is no longer valid? We recall that, assuming $z \gg h$, a collective approach has already been investigated [89].

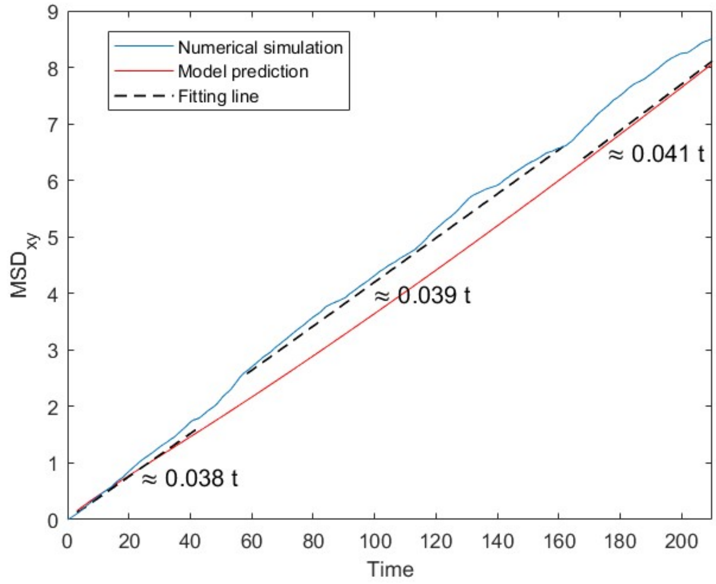
2. What is the critical number density n_c above which our model is no longer valid? In other words, can we quantify the assumption of "low number density", which was also made in [152, 117]?

While we do not know how to answer the first question, we propose a possible criterion for roughly estimating the range of number densities n to which our model applies. Consider a tracer at height z_1 and denote by $\mathbb{E}[r]$ the average distance of its projection onto the plane $\{z = 0\}$ to its nearest neighbour. By (7.4.22), it holds that $\mathbb{E}[r] = 1/(2\sqrt{n})$. Let $\max\{[\|\Delta_{xy}(z_1)\|\}$ denote the maximum norm of the displacement, in

[19] To be more precise, the Dataset S01 we use corresponds to the case considered in Movie S16; see [155, Supporting Information] for more details.

the xy variables only, of the tracer assuming an actuator acts on it for a time $\bar{t} = \int_0^{+\infty} t p_t(t) dt$. Then, a possible criterion might be [20]:

$$\max\{\|\Delta_{xy}(z_1)\|\} < \mathbb{E}[r] \implies n < \frac{1}{4\max\{\|\Delta_{xy}(z_1)\|\}^2}. \tag{7.4.23}$$



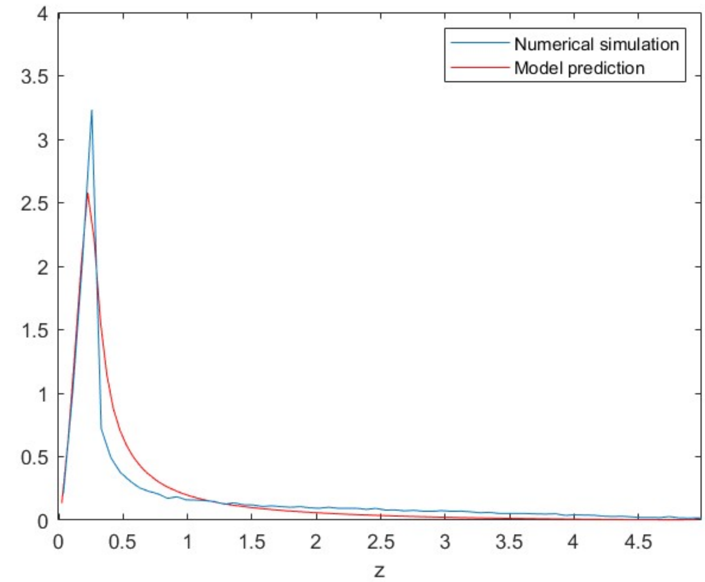


Figure 7.17: (Up) The mean squared displacement in the xy variables (MSD_{xy}) up to time $T = 210\,\mathrm{s}$ considering a number density $n = 0.005$, the empirical probability distribution p_t for activation times and an initial height $z_1 = 0.25$. We compare the model prediction (in red) with the MSD_{xy} obtained by randomly selecting 10 actuators in a circle of radius $R = 25$ and center $C = (45, 35)$, and a random initial position (in the xy variables) for the tracer inside a circle of center C and radius $r = 15$. The mean squared displacement is computed over 5×10^4 realisations. The dotted black lines give the slopes of the curves. (Down) The height distribution at time $T = 210\,\mathrm{s}$ given by the model prediction (in red) and by the numerical experiments described above (in blue).

[20] We remark that, in our numerical simulations, $\bar{t} \approx 7.6\,\mathrm{s}$ and the maximum norms of displacements for a tracer at heights 0.25 and 0.75, considering $L_c = 200\,\mu\mathrm{m}$ as the unit of measurement, are of 3.4 and 4.2. So, the right hand side of (7.4.23) is equal to 0.02 and 0.014 respectively, which are bigger than the number density $n = 0.005$ when considering only 10 actuators. Instead, the inequality does not hold when considering 40 actuators and a number density $n = 0.02$, and indeed Figure 7.19 shows that the model fails in predicting the mean squared displacement.

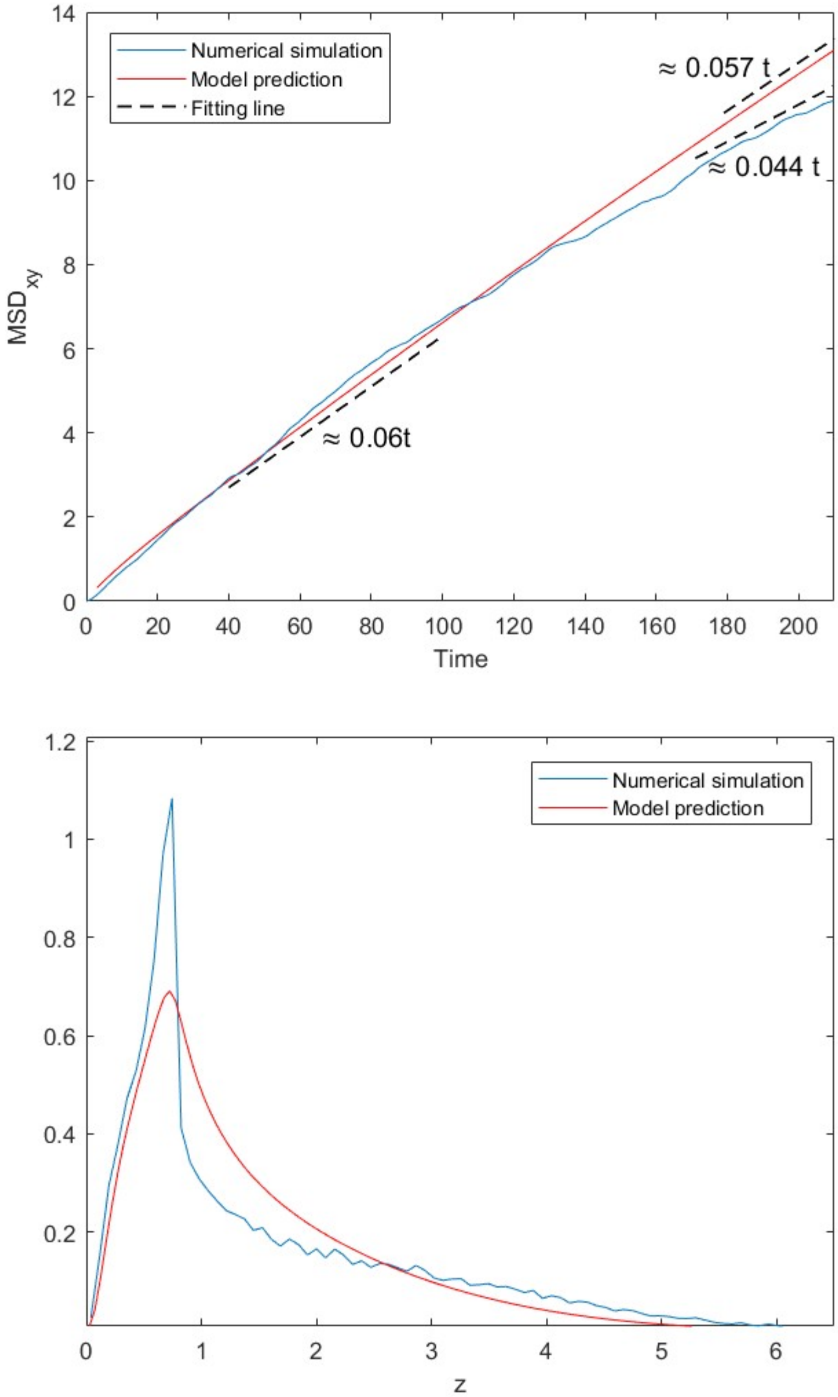


Figure 7.18: (Up) The mean squared displacement in the xy variables (MSD_{xy}) up to time $T = 210\,\mathrm{s}$ considering a number density $n = 0.005$, the empirical probability distribution p_t for activation times and an initial height $z_1 = 0.75$. We compare the model prediction (in red) with the MSD_{xy} obtained by randomly selecting 10 actuators in a circle of radius $R = 25$ and center $C = (45, 35)$, and a random initial position (in the xy variables) for the tracer inside a circle of center C and radius $r = 15$. The mean squared displacement is computed over 5×10^4 realisations. The dotted black lines give the slopes of the curves. (Down) The height distribution at time $T = 210\,\mathrm{s}$ given by the model prediction (in red) and by the numerical experiments described above (in blue).

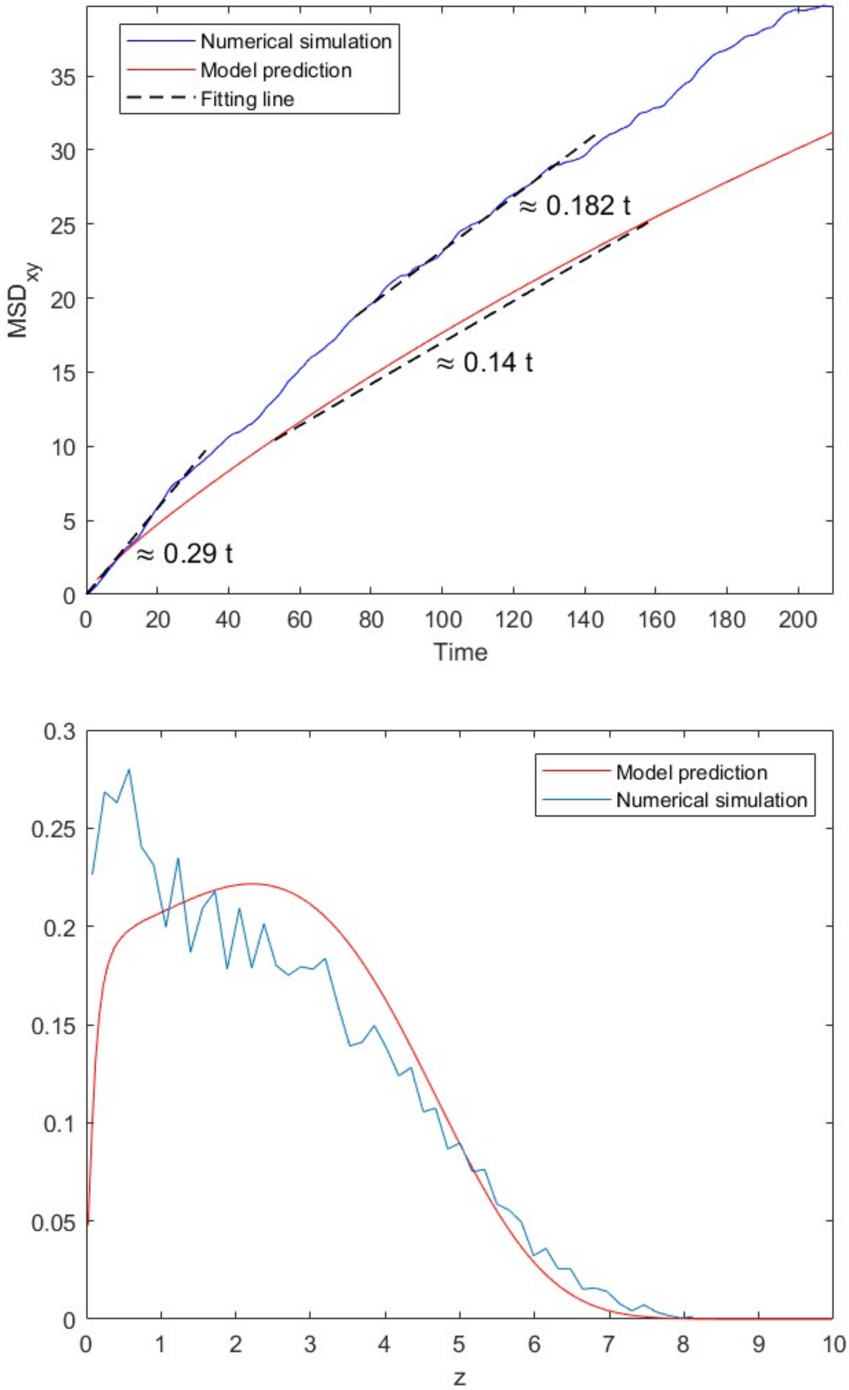


Figure 7.19: (Up) The mean squared displacement in the xy variables (MSD_{xy}) up to time $T = 210\,\mathrm{s}$ considering a number density $n = 0.02$, the empirical probability distribution p_t for activation times and an initial height $z_1 = 0.75$. We compare the model prediction (in red) with the MSD_{xy} obtained by randomly selecting 40 actuators in a circle of radius $R = 25$ and center $C = (45, 35)$, and a random initial position (in the xy variables) for the tracer inside a circle of center C and radius $r = 15$. The mean squared displacement is computed over 10^4 realisations. The dotted black lines give the slopes of the curves. (Down) The height distribution at time $T = 210\,\mathrm{s}$ given by the model prediction (in red) and by the numerical experiments described above (in blue).

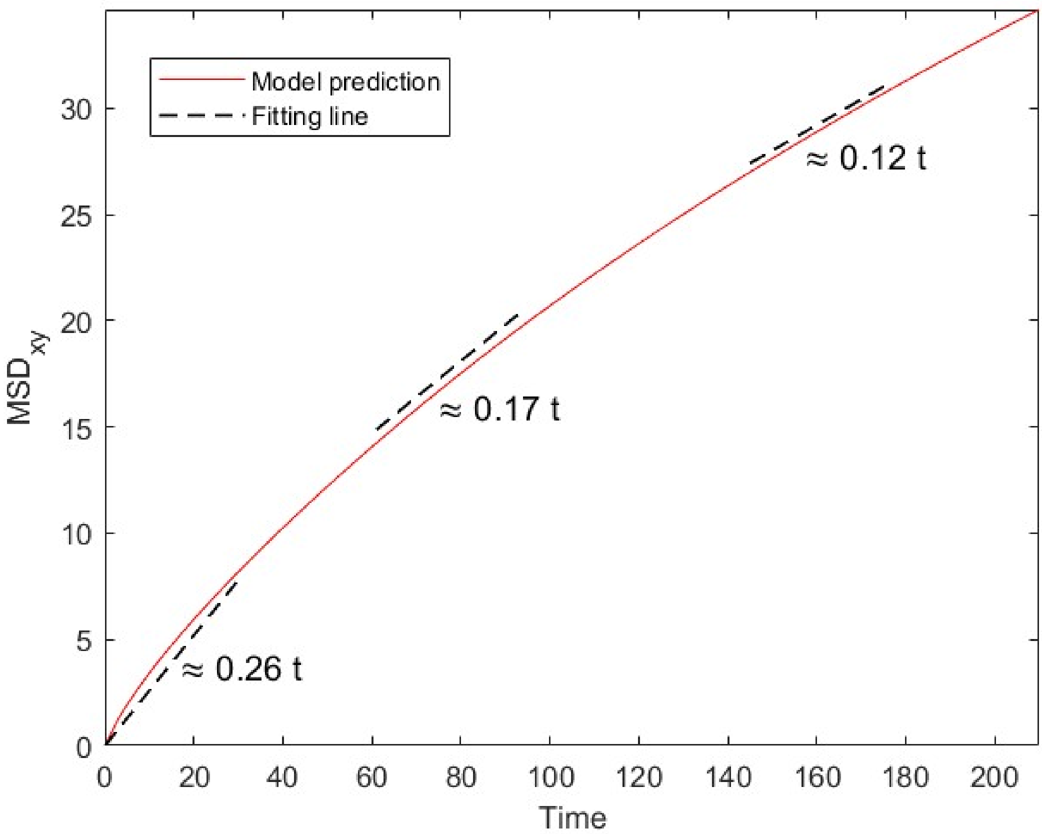


Figure 7.20: Model prediction of the mean squared displacement MSD_{xy} up to time $T = 210\,\mathrm{s}$ considering a number density $n = 0.03$, the empirical probability p_t for activation times and an initial height $z_1 = 0.75$. The dotted black lines give the slopes of the curve.

Bibliography

[1] Robert A. Adams. *Sobolev spaces*, volume Vol. 65 of *Pure and Applied Mathematics*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[2] Josephine Ainley, Sandra Durkin, Rafael Embid, Priya Boindala, and Ricardo Cortez. The method of images for regularized Stokeslets. *Journal of Computational Physics*, 227(9):4600–4616, 2008.

[3] Michael Aizenman. On vector fields as generators of flows: a counterexample to Nelson's conjecture. *Ann. of Math. (2)*, 107(2):287–296, 1978.

[4] Giovanni Alberti. Generalized N-property and Sard theorem for Sobolev maps. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 23(4):477–491, 2012.

[5] Giovanni Alberti, Stefano Bianchini, and Gianluca Crippa. A uniqueness result for the continuity equation in two dimensions. *J. Eur. Math. Soc. (JEMS)*, 16(2):201–234, 2014.

[6] Giovanni Alberti, Tommaso Cortopassi, and Antonio De Simon. Stirring by unicellular feeders. (in preparation).

[7] Luigi Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, 158(2):227–260, 2004.

[8] Luigi Ambrosio. Transport equation and Cauchy problem for non-smooth vector fields. In *Calculus of variations and nonlinear partial differential equations*, volume 1927 of *Lecture Notes in Math.*, pages 1–41. Springer, Berlin, 2008.

[9] Luigi Ambrosio and Patrick Bernard. Uniqueness of signed measures solving the continuity equation for Osgood vector fields. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 19(3):237–245, 2008.

[10] Luigi Ambrosio and Gianluca Crippa. Continuity equations and ODE flows with non-smooth velocity. *Proc. Roy. Soc. Edinburgh Sect. A*, 144(6):1191–1244, 2014.

[11] Luigi Ambrosio, Camillo De Lellis, and Jan Malý. On the chain rule for the divergence of BV-like vector fields: applications, partial results, open problems. In *Perspectives in nonlinear partial differential equations*, volume 446 of *Contemp. Math.*, pages 31–67. Amer. Math. Soc., Providence, RI, 2007.

[12] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[13] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[14] Luigi Ambrosio, Myriam Lecumberry, and Stefania Maniglia. Lipschitz regularity and approximate differentiability of the Diperna-Lions flow. *Rendiconti del Seminario Matematico della Università di Padova*, 114:29–50, 2005.

[15] Paolo Baldi. *Stochastic calculus*. Universitext. Springer, Cham, 2017. An introduction through theory and exercises.

[16] L. Batteux, T. Gallouët, R. Herbin, J. C. Latché, and P. Poullet. Convergence of the MAC scheme for the incompressible Navier-Stokes equations with variable density and viscosity. *Math. Comp.*, 92(342):1595–1631, 2023.

[17] Anatole Beck. Uniqueness of flow solutions of differential equations. In *Recent advances in topological dynamics (Proc. Conf. Topological Dynamics, Yale Univ., New Haven, Conn., 1972; in honor of Gustav Arnold Hedlund)*, volume Vol. 318 of *Lecture Notes in Math.*, pages 30–50. Springer, Berlin-New York, 1973.

[18] Fethi Ben Belgacem and Pierre-Emmanuel Jabin. Compactness for nonlinear continuity equations. *J. Funct. Anal.*, 264(1):139–168, 2013.

[19] F. Ben Belgacem and P.-E. Jabin. Convergence of numerical approximations to non-linear continuity equations with rough force fields. *Arch. Ration. Mech. Anal.*, 234(2):509–547, 2019.

[20] Howard C Berg. *Random walks in biology*. Princeton University Press, 1993.

[21] Stefano Bianchini and Paolo Bonicatto. A uniqueness result for the decomposition of vector fields in $\mathbb{R}^d$. *Invent. Math.*, 220(1):255–393, 2020.

[22] J. R. Blake. A spherical envelope approach to ciliary propulsion. *Journal of Fluid Mechanics*, 46(1):199–208, 1971.

[23] John R Blake. A note on the image system for a stokeslet in a no-slip boundary. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 70(2), pages 303–310. Cambridge University Press, 1971.

[24] John R Blake and Allen T Chwang. Fundamental singularities of viscous flow: part i: the image systems in the vicinity of a stationary no-slip boundary. *Journal of Engineering Mathematics*, 8(1):23–29, 1974.

[25] Paolo Bonicatto, Gennaro Ciampa, and Gianluca Crippa. On the advection-diffusion equation with rough coefficients: weak solutions and vanishing viscosity. *J. Math. Pures Appl. (9)*, 167:204–224, 2022.

[26] Paolo Bonicatto, Gennaro Ciampa, and Gianluca Crippa. Weak and parabolic solutions of advection-diffusion equations with rough velocity field. *J. Evol. Equ.*, 24(1):Paper No. 1, 16, 2024.

[27] Paolo Bonicatto, Giacomo Del Nin, and Filip Rindler. Existence and uniqueness for the transport of currents by Lipschitz vector fields. *J. Funct. Anal.*, 286(7):Paper No. 110315, 24, 2024.

[28] Paolo Bonicatto and Nikolay A. Gusev. Non-uniqueness of signed measure-valued solutions to the continuity equation in presence of a unique flow. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 30(3):511–531, 2019.

[29] François Bouchut. Renormalized solutions to the Vlasov equation with coefficients of bounded variation. *Arch. Ration. Mech. Anal.*, 157(1):75–90, 2001.

[30] François Bouchut and Gianluca Crippa. Lagrangian flows for vector fields with gradient given by a singular integral. *J. Hyperbolic Differ. Equ.*, 10(2):235–282, 2013.

[31] Franck Boyer. Analysis of the upwind finite volume method for general initial- and boundary-value transport problems. *IMA J. Numer. Anal.*, 32(4):1404–1439, 2012.

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

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

[34] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[35] Robert Brown. XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The philosophical magazine*, 4(21):161–173, 1828.

[36] Elia Bruè, Maria Colombo, and Camillo De Lellis. Positive solutions of transport equations and classical nonuniqueness of characteristic curves. *Arch. Ration. Mech. Anal.*, 240(2):1055–1090, 2021.

[37] Elia Bruè, Maria Colombo, and Anuj Kumar. Sharp Nonuniqueness in the Transport Equation with Sobolev Velocity Field, 2024. arXiv:2405.01670.

[38] Tristan Buckmaster. Onsager's conjecture almost everywhere in time. *Comm. Math. Phys.*, 333(3):1175–1198, 2015.

[39] Tristan Buckmaster, Camillo De Lellis, Philip Isett, and László Székelyhidi, Jr. Anomalous dissipation for $1/5$-Hölder Euler flows. *Ann. of Math. (2)*, 182(1):127–172, 2015.

[40] Tristan Buckmaster and Vlad Vicol. Nonuniqueness of weak solutions to the Navier-Stokes equation. *Ann. of Math. (2)*, 189(1):101–144, 2019.

[41] Tristan Buckmaster and Vlad Vicol. Convex integration constructions in hydrodynamics. *Bull. Amer. Math. Soc. (N.S.)*, 58(1):1–44, 2021.

[42] Martin Burger, Yasmin Dolak-Struss, and Christian Schmeiser. Asymptotic analysis of an advection-dominated chemotaxis model in multiple spatial dimensions. *Commun. Math. Sci.*, 6(1):1–28, 2008.

[43] Laura Caravenna and Gianluca Crippa. A directional Lipschitz extension lemma, with applications to uniqueness and Lagrangianity for the continuity equation. *Comm. Partial Differential Equations*, 46(8):1488–1520, 2021.

[44] Nicolas Champagnat and Pierre-Emmanuel Jabin. Well posedness in any dimension for Hamiltonian flows with non BV force terms. *Comm. Partial Differential Equations*, 35(5):786–816, 2010.

[45] Nicolas Champagnat and Pierre-Emmanuel Jabin. Strong solutions to stochastic differential equations with rough coefficients. *Ann. Probab.*, 46(3):1498–1541, 2018.

[46] Gennaro Ciampa, Tommaso Cortopassi, Gianluca Crippa, Raffaele D'Ambrosio, and Stefano Spirito. A priori error estimates for the θ-method for the flow of nonsmooth velocity fields, 2025. arXiv:2506.02747 (to appear in SIAM Journal on Numerical Analysis).

[47] Earl A. Coddington and Norman Levinson. *Theory of ordinary differential equations*. McGraw-Hill Book Co., Inc., New York-Toronto-London, 1955.

[48] Edward A Codling, Michael J Plank, and Simon Benhamou. Random walk models in biology. *Journal of the Royal society interface*, 5(25):813–834, 2008.

[49] F. Colombini, T. Luo, and J. Rauch. Uniqueness and nonuniqueness for nonsmooth divergence free transport. In *Seminaire: Équations aux Dérivées Partielles, 2002–2003*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XXII, 21. École Polytech., Palaiseau, 2003.

[50] Ferruccio Colombini and Nicolas Lerner. Uniqueness of L^∞ solutions for a class of conormal BV vector fields. In *Geometric analysis of PDE and several complex variables*, volume 368 of *Contemp. Math.*, pages 133–156. Amer. Math. Soc., Providence, RI, 2005.

[51] Maria Colombo, Gianluca Crippa, and Stefano Spirito. Renormalized solutions to the continuity equation with an integrable damping term. *Calc. Var. Partial Differential Equations*, 54(2):1831–1845, 2015.

[52] S. Corrsin. Turbulent flow. *American Scientist*, 49(3):300–325, 1961.

[53] Ricardo Cortez. The method of regularized Stokeslets. *SIAM Journal on Scientific Computing*, 23(4):1204–1225, 2001.

[54] Tommaso Cortopassi. A note on a L^p stability estimate for regular Lagrangian flows, 2023. arXiv:2310.03871.

[55] Tommaso Cortopassi. A current based approach for the uniqueness of the continuity equation, 2024. arXiv:2402.10719 (submitted).

[56] Tommaso Cortopassi. An explicit Euler method for Sobolev vector fields with applications to the continuity equation on non Cartesian grids. *J. Math. Pures Appl. (9)*, 199:Paper No. 103722, 35, 2025.

[57] Gianluca Crippa. *The flow associated to weakly differentiable vector fields*, volume 12 of *Tesi. Scuola Normale Superiore di Pisa (Nuova Series) [Theses of Scuola Normale Superiore di Pisa (New Series)]*. Edizioni della Normale, Pisa, 2009.

[58] Gianluca Crippa and Camillo De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.

[59] Sara Daneri and László Székelyhidi, Jr. Non-uniqueness and h-principle for Hölder-continuous weak solutions of the Euler equations. *Arch. Ration. Mech. Anal.*, 224(2):471–514, 2017.

[60] Charles Darwin. Note on hydrodynamics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 49(2):342–354, 1953.

[61] Camillo De Lellis. Notes on hyperbolic systems of conservation laws and transport equations. In *Handbook of differential equations: evolutionary equations. Vol. III*, Handb. Differ. Equ., pages 277–382. Elsevier/North-Holland, Amsterdam, 2007.

[62] Camillo De Lellis, Piotr Gwiazda, and Agnieszka Świerczewska-Gwiazda. Transport equations with integral terms: existence, uniqueness and stability. *Calc. Var. Partial Differential Equations*, 55(5):Art. 128, 17, 2016.

[63] Camillo De Lellis and László Székelyhidi, Jr. The Euler equations as a differential inclusion. *Ann. of Math. (2)*, 170(3):1417–1436, 2009.

[64] Camillo De Lellis and László Székelyhidi, Jr. Dissipative continuous Euler flows. *Invent. Math.*, 193(2):377–407, 2013.

[65] JMPQ Delgado. Molecular diffusion coefficients of organic compounds in water at different temperatures. *Journal of phase Equilibria and Diffusion*, 28(5):427–432, 2007.

[66] Nicolas Depauw. Non unicité des solutions bornées pour un champ de vecteurs BV en dehors d'un hyperplan. *C. R. Math. Acad. Sci. Paris*, 337(4):249–252, 2003.

[67] Bruno Despres. Lax theorem and finite volume schemes. *Math. Comp.*, 73(247):1203–1234, 2004.

[68] W. K. Dewar, R. J. Bingham, R. L. Iverson, D. P. Nowacek, L. C. St Laurent, and P. H. Wiebe. Did marine life cause ocean mixing? *Journal of Marine Research*, 64(4):541–561, 2006.

[69] Eleonora Di Nezza, Giampiero Palatucci, and Enrico Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[70] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.

[71] I Eames, SE Belcher, and JCR1297867 Hunt. Drift, partial drift and Darwin's proposition. *Journal of Fluid Mechanics*, 275:201–223, 1994.

[72] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4, 1905.

[73] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[74] R. Eymard, T. Gallouët, R. Herbin, and J. C. Latché. A convergent finite element-finite volume scheme for the compressible Stokes problem. II. The isentropic case. *Math. Comp.*, 79(270):649–675, 2010.

[75] Herbert Federer. *Geometric measure theory*, volume Band 153 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag New York, Inc., New York, 1969.

[76] Adolf Fick. Ueber diffusion. *Annalen der physik*, 170(1):59–86, 1855.

[77] Alessio Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.

[78] F. Flandoli, M. Gubinelli, and E. Priola. Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.*, 180(1):1–53, 2010.

[79] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields*, 162(3-4):707–738, 2015.

[80] T. Gallouët, R. Herbin, and J.-C. Latché. A convergent finite element-finite volume scheme for the compressible Stokes problem. I. The isothermal case. *Math. Comp.*, 78(267):1333–1352, 2009.

[81] T. Gallouët, R. Herbin, and J.-C. Latché. Kinetic energy control in explicit finite volume discretizations of the incompressible and compressible Navier-Stokes equations. *Int. J. Finite Vol.*, 7(2):6, 2010.

[82] Thierry Gallouët, Laura Gastaldo, Raphaele Herbin, and Jean-Claude Latché. An unconditionally stable pressure correction scheme for the compressible barotropic Navier-Stokes equations. *M2AN Math. Model. Numer. Anal.*, 42(2):303–331, 2008.

[83] Michael C Gregg, Thomas B Sanford, and David P Winkel. Mixing by swimming organisms. *Science*, 316(5826):838, 2007.

[84] T. H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Ann. of Math. (2)*, 20(4):292–296, 1919.

[85] A. Grothendieck. La théorie de Fredholm. *Bull. Soc. Math. France*, 84:319–384, 1956.

[86] Morton E. Gurtin. *An introduction to continuum mechanics*, volume 158 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981.

[87] Nikolay A. Gusev. A necessary and sufficient condition for existence of measurable flow of a bounded Borel vector field. *Mosc. Math. J.*, 18(1):85–92, 2018.

[88] Nikolay A. Gusev. On the one-dimensional continuity equation with a nearly incompressible vector field. *Commun. Pure Appl. Anal.*, 18(2):559–568, 2019.

[89] Francisca Guzmán-Lastra, Hartmut Löwen, and Arnold JTM Mathijssen. Active carpets drive non-equilibrium diffusion and enhanced molecular fluxes. *Nature communications*, 12(1):1906, 2021.

[90] P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980.

[91] Peter Hänggi and Fabio Marchesoni. Introduction: 100years of Brownian motion. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15(2), 2005.

[92] Philip Hartman. *Ordinary differential equations*, volume 38 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Corrected reprint of the second (1982) edition [Birkhäuser, Boston, MA; MR0658490 (83e:34002)], With a foreword by Peter Bates.

[93] Skeeve (https://mathoverflow.net/users/44463/skeeve). Ode with a measurable vector field. MathOverflow. URL:https://mathoverflow.net/q/323449 (version: 2019-02-17).

[94] Sergio Idelsohn, Norberto Nigro, Alejandro Limache, and Eugenio Oñate. Large time-step explicit integration method for solving problems with dominant convection. *Computer Methods in Applied Mechanics and Engineering*, 217-220:168–185, 2012.

[95] Philip Isett. A proof of Onsager's conjecture. *Ann. of Math. (2)*, 188(3):871–963, 2018.

[96] Pierre-Emmanuel Jabin and Datong Zhou. Discretizing advection equations with rough velocity fields on non-Cartesian grids. *Quart. Appl. Math.*, 82(2):229–303, 2024.

[97] Benjamin Jourdain and Stéphane Menozzi. Convergence rate of the Euler-Maruyama scheme applied to diffusion processes with $L^q - L^\rho$ drift coefficient and additive noise. *Ann. Appl. Probab.*, 34(1B):1663–1697, 2024.

[98] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[99] Kakani Katija and John O Dabiri. A viscosity-enhanced mechanism for biogenic mixing and transport in the ocean. *Nature*, 460:624–626, 2009.

[100] Sangtae Kim and Seppo J Karrila. *Microhydrodynamics: principles and selected applications*. Butterworth-Heinemann, 2013.

[101] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[102] Anuj Kumar. Nonuniqueness of trajectories on a set of full measure for Sobolev vector fields. *Arch. Ration. Mech. Anal.*, 248(6):Paper No. 114, 26, 2024.

[103] H. Kunita. Stochastic differential equations and stochastic flows of diffeomorphisms. In *École d'été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 143–303. Springer, Berlin, 1984.

[104] N. N. Kuznecov. The accuracy of certain approximate methods for the computation of weak solutions of a first order quasilinear equation. *Ž. Vyčisl. Mat i Mat. Fiz.*, 16(6):1489–1502, 1627, 1976.

[105] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Uraltseva. *Linear and quasilinear equations of parabolic type*, volume Vol. 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1968. Translated from the Russian by S. Smith.

[106] Lev Davidovich Landau and Evgeny Mikhailovich Lifshitz. *Fluid Mechanics: Volume 6*, volume 6. Elsevier, 1987.

[107] Eric Lauga. *The fluid dynamics of cell motility*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2020.

[108] Peter D. Lax. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Comm. Pure Appl. Math.*, 7:159–193, 1954.

[109] C. Le Bris and P.-L. Lions. Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications. *Ann. Mat. Pura Appl. (4)*, 183(1):97–130, 2004.

[110] 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(7-9):1272–1317, 2008.

[111] Claude Le Bris and Pierre-Louis Lions. *Parabolic equations with irregular data and related issues—applications to stochastic differential equations*, volume 4 of *De Gruyter Series in Applied and Numerical Mathematics*. De Gruyter, Berlin, [2019] ©2019.

[112] Henri Lebesgue. Intégrale, longueur, aire. *Annali di Matematica Pura ed Applicata (1898-1922)*, 7(1):231–359, 1902.

[113] Randall J. LeVeque. *Numerical methods for conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 1992.

[114] Randall J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.

[115] Huaiqian Li and Dejun Luo. Quantitative stability estimates for Fokker-Planck equations. *J. Math. Pures Appl. (9)*, 122:125–163, 2019.

[116] M J Lighthill. On the squirming motion of nearly spherical deformable bodies through liquids at very small Reynolds numbers. *Communications on Pure and Applied Mathematics*, 5(2):109–118, 1952.

[117] Zhi Lin, Jean-Luc Thiffeault, and Stephen Childress. Stirring by squirmers. *J. Fluid Mech.*, 669:167–177, 2011.

[118] Ernest Lindelöf. Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 116(3):454–457, 1894.

[119] Nadav Liron and S Mochon. Stokes flow for a stokeslet between two parallel flat plates. *Journal of Engineering Mathematics*, 10(4):287–303, 1976.

[120] De Jun Luo. Fokker-Planck type equations with Sobolev diffusion coefficients and BV drift coefficients. *Acta Math. Sin. (Engl. Ser.)*, 29(2):303–314, 2013.

[121] Dejun Luo. Uniqueness of degenerate Fokker-Planck equations with weakly differentiable drift whose gradient is given by a singular integral. *Electron. Commun. Probab.*, 19:no. 43, 14, 2014.

[122] Jan R Magnus and Heinz Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. John Wiley & Sons, 2019.

[123] Benoit Merlet. L^∞- and L^2-error estimates for a finite volume approximation of linear advection. *SIAM J. Numer. Anal.*, 46(1):124–150, 2007/08.

[124] Benoît Merlet and Julien Vovelle. Error estimate for finite volume scheme. *Numer. Math.*, 106(1):129–155, 2007.

[125] Stefano Modena and Gabriel Sattig. Convex integration solutions to the transport equation with full dimensional concentration. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 37(5):1075–1108, 2020.

[126] Stefano Modena and László Székelyhidi, Jr. Non-uniqueness for the transport equation with Sobolev vector fields. *Ann. PDE*, 4(2):Paper No. 18, 38, 2018.

[127] Stefano Modena and László Székelyhidi, Jr. Non-renormalized solutions to the continuity equation. *Calc. Var. Partial Differential Equations*, 58(6):Paper No. 208, 30, 2019.

[128] Víctor Navarro-Fernández and André Schlichting. Error estimates for a finite volume scheme for advection-diffusion equations with rough coefficients. *ESAIM Math. Model. Numer. Anal.*, 57(4):2131–2158, 2023.

[129] Víctor Navarro-Fernández, André Schlichting, and Christian Seis. Optimal stability estimates and a new uniqueness result for advection-diffusion equations. *Pure Appl. Anal.*, 4(3):571–596, 2022.

[130] Quoc-Hung Nguyen. Quantitative estimates for regular Lagrangian flows with BV vector fields. *Comm. Pure Appl. Math.*, 74(6):1129–1192, 2021.

[131] Bernt Øksendal. *Stochastic differential equations: An introduction with applications*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003.

[132] Lars Onsager. Statistical hydrodynamics. *Il Nuovo Cimento (1943-1954)*, 6(Suppl 2):279–287, 1949.

[133] W. F. Osgood. Beweis der Existenz einer Lösung der Differentialgleichung $\frac{dy}{dx} = f(x, y)$ ohne Hinzunahme der Cauchy-Lipschitz'schen Bedingung. *Monatsh. Math. Phys.*, 9(1):331–345, 1898.

[134] G. P. Patil and C. R. Rao. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, 34(2):179–189, 1978.

[135] G. Peano. Démonstration de l'intégrabilité des équations différentielles ordinaires. *Math. Ann.*, 37(2):182–228, 1890.

[136] J. B. Perrin. Mouvement brownien et réalité moléculaire. *Ann. de Chimie et de Physique (VIII)*, 18, 1909.

[137] Jules Pitcho and Massimo Sorella. Almost everywhere nonuniqueness of integral curves for divergence-free Sobolev vector fields. *SIAM J. Math. Anal.*, 55(5):4640–4663, 2023.

[138] Edward M Purcell. Life at low Reynolds number. *American Journal of Physics*, 45(1):3–11, 1977.

[139] Daniel Revuz. *Markov chains*, volume 11. Elsevier, 2008.

[140] Michael Röckner and Xicheng Zhang. Weak uniqueness of Fokker-Planck equations with degenerate and bounded coefficients. *C. R. Math. Acad. Sci. Paris*, 348(7-8):435–438, 2010.

[141] Walter Rudin. Well-distributed measurable sets. *The American Mathematical Monthly*, 90(1):41–42, 1983.

[142] André Schlichting and Christian Seis. Convergence rates for upwind schemes with rough coefficients. *SIAM J. Numer. Anal.*, 55(2):812–840, 2017.

[143] André Schlichting and Christian Seis. Analysis of the implicit upwind finite volume scheme with rough coefficients. *Numer. Math.*, 139(1):155–186, 2018.

[144] Christian Seis. A quantitative theory for the continuity equation. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 34(7):1837–1850, 2017.

[145] Christian Seis. Optimal stability estimates for continuity equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 148(6):1279–1296, 2018.

[146] Alexander Shaposhnikov and Lukas Wresch. Pathwise vs. path-by-path uniqueness. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(3):1640–1649, 2022.

[147] Boris I Shraiman and Eric D Siggia. Scalar turbulence. *Nature*, 405(6787):639–646, 2000.

[148] Leon Simon. *Lectures on geometric measure theory*, volume 3 of *Proceedings of the Centre for Mathematical Analysis, Australian National University*. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[149] Elias M. Stein. *Singular integrals and differentiability properties of functions*, volume No. 30 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1970.

[150] Roman Stocker. Marine microbes see a sea of gradients. *Science*, 338(6107):628–633, 2012.

[151] John R Taylor and Roman Stocker. Trade-offs of chemotactic foraging in turbulent water. *Science*, 338(6107):675–679, 2012.

[152] Jean-Luc Thiffeault and Stephen Childress. Stirring by swimming bodies. *Physics Letters A*, 374(34):3487–3490, 2010.

[153] Chad M. Topaz and Andrea L. Bertozzi. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM J. Appl. Math.*, 65(1):152–174, 2004.

[154] Dario Trevisan. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.*, 21:Paper No. 22, 41, 2016.

[155] Régis Turuban, Giovanni Noselli, Alfred Beran, and Antonio DeSimone. Cooperative mixing through hydrodynamic interactions in stylonychia lemnae. *Proceedings of the National Academy of Sciences*, 122(37):e2500588122, 2025.

[156] A. Ju. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S.)*, 111(153)(3):434–452, 480, 1980.

[157] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[158] Andre W Visser. Biogenic mixing in the ocean: effects of large animals. *Limnology and Oceanography*, 52(1):447–460, 2007.

[159] Noel J. Walkington. Convergence of the discontinuous Galerkin method for discontinuous solutions. *SIAM J. Numer. Anal.*, 42(5):1801–1817, 2005.

[160] Chia-Shun Yih. Evolution of Darwinian drift. *Journal of Fluid Mechanics*, 347:1–11, 1997.