University of Padua

Department of Mathematics "Tullio Levi-Civita"

Osgood meets DiPerna-Lions

Leonardo Franchi

Student ID: 2107788

Master's Degree in Mathematics

Supervisor: Prof. Guido De Philippis

Academic Year 2024/2025

Contents

Introduction

The aim of this thesis is to study the following three equations:

Ordinary Differential Equation (ODE)

$$\dot{\gamma}(t) = b(t, \gamma(t)) \tag{1}$$

Continuity Equation

$$\partial_t u + \nabla \cdot (bu) = 0 \tag{2}$$

Transport Equation

$$\partial_t u + b \cdot \nabla u = 0 \tag{3}$$

Here, $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a vector field and $\nabla\cdot$ is the divergence operator. Depending on the regularity of the vector field b, we will give a precise meaning to these equations (either classical or in the distributional sense) and study the main properties of their solutions: existence, uniqueness, and regularity.

In **Chapter 1**, we will present classical results when the vector field is sufficiently smooth, meaning at least locally Lipschitz.

In **Chapter 2**, we will introduce the key ideas developed by DiPerna and Lions in their seminal work [**DL89**]. Specifically, we will prove that, in a certain sense, existence and uniqueness hold for the continuity equation by introducing the key notion of renormalized solutions. In the second part of this chapter, we will build on these ideas, using the superposition principle and the methods introduced by Ambrosio in [**Amb04**], to extend these results to the ODE in a low-regularity setting.

In **Chapter 3**, we will explore a different approach to this problem, focusing on deriving *a priori* estimates directly at the PDE level. This method was first introduced by Crippa and De Lellis in [**CL08**]. Additionally, this chapter will include a discussion on the optimal regularity of solutions to the continuity equation, following the work of Brue and Nguyen in [**BN21**].

In **Chapter 4**, we generalize the results of the previous two chapters under weaker assumptions. We consider functions that, in a certain sense, have slightly less than one derivative in L^p, where this notion is determined by an Osgood function. In the first part, we show that under these assumptions a Regular Lagrangian Flow always exists. In the second part, we prove that, in a certain regime defined by the Osgood function, solutions of the transport equation are renormalized. This result extends the renormalization property of solutions beyond the classical Sobolev threshold. The key idea is to use the Littlewood–Paley decomposition to first prove a regularity result for the solutions of the transport equation, and then combine this regularity with that of the vector field to establish the renormalization property via a commutator estimate. This latter result is joint work with Guido De Philippis.

Ringraziamenti

Innanzitutto, un sentito ringraziamento va al mio relatore, Guido De Philippis, per l'enorme supporto e il tempo che mi ha dedicato. Grazie a lui ho imparato moltissimo e sono riuscito a raggiungere obiettivi che, solo un anno fa, mi sembravano lontanissimi. Queste poche frasi non riusciranno a riassumere quanto gli sia riconoscente.

Vorrei sinceramente ringraziare anche i professori Paolo Ciatti e Davide Vittone per il costante supporto e per le numerose conversazioni che abbiamo avuto, sempre stimolanti e preziose.

Infine, non per importanza, vorrei ringraziare la mia famiglia e tutte le altre persone che mi sono state vicine al di fuori dell'ambito accademico. Non le citerò una ad una qui, ma sanno chi sono e spero di riuscire a ringraziarle ogni giorno per il loro supporto e per tutte le belle cose vissute insieme.

CHAPTER 1

Classical Theory in the Smooth Framework

1. ODE in the smooth framework

In this section, we study the ordinary differential equation (1) in the case where the vector field is regular enough. We begin by specifying what we mean by a solution in this context.

Let $b : \Omega \subset \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous vector field, defined on an open set Ω. A *(classical) solution* of the ordinary differential equation

$$\dot{\gamma}(t) = b(t, \gamma(t)) \tag{4}$$

is a function $\gamma \in C^1([t_1, t_2]; \mathbb{R}^d)$, defined on an interval $[t_1, t_2] \subset \mathbb{R}$, such that $(t, \gamma(t)) \in \Omega$ for every $t \in [t_1, t_2]$ and such that the equation (4) holds pointwise. In this case, we also refer to γ as an *integral curve* or a *characteristic curve* of the vector field b.

Given an initial point $(t_0, x_0) \in \Omega$, we consider the Cauchy problem:

$$\begin{cases} \dot{\gamma}(t) = b(t, \gamma(t)) \\ \gamma(t_0) = x_0. \end{cases} \tag{5}$$

A solution to this problem is a function $\gamma \in C^1([t_1, t_2]; \mathbb{R}^d)$ that satisfies the differential equation (4) and the initial condition $\gamma(t_0) = x_0$, with $t_0 \in [t_1, t_2]$ and $(t, \gamma(t)) \in \Omega$ for all t in the interval.

In this setting, the equation can be reformulated as the problem of finding a function γ such that, for every $t \in [t_1, t_2]$, the following integral equation holds:

$$\gamma(t) = x_0 + \int_{t_0}^{t} b(s, \gamma(s))\, ds. \tag{6}$$

This equation can be further reformulated as the problem of finding a fixed point of the operator

$$T : C^0 \to C^0, \quad T(f)(t) := x_0 + \int_{t_0}^{t} b(s, f(s))\, ds.$$

This formulation allows us to prove our first existence result: *Peano's Theorem.*

Theorem 1.1 (Peano) Let $b : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function. Then, for any initial condition $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^d$ the ordinary differential equation

$$\begin{cases} \dot{\gamma}(t) = b(t, \gamma(t)) \\ \gamma(t_0) = x_0 \end{cases}$$

admits at least one solution $\gamma \in C^1([t_0 - \delta, t_0 + \delta], \mathbb{R}^d)$ for some fixed δ that depends on b, t_0 and x_0.

Proof. We want to apply the Schauder fixed-point theorem, and we will work using the supremum norm. Let $\alpha, \beta > 0$ be fixed, and consider the set

$$X := \{(t, x) \in \mathbb{R} \times \mathbb{R}^d : |t - t_0| \leq \alpha, |x - x_0| \leq \beta\}.$$

Let M denote the maximum of b on X; we will show that any $\delta < \min\left\{\alpha, \frac{\beta}{M}\right\}$ works.

First, note that for a solution of (5), we have

$$|\gamma(t) - x_0| = \left| \int_0^t b(s, \gamma(s))\, ds \right| \leq |t|M \leq |\delta|M \leq \beta.$$

Thus, it is natural to consider the space of continuous functions Y, defined as follows:

$$Y := \left\{ \gamma \in C^0([t_0-\delta, t_0+\delta], \mathbb{R}^d) : \gamma(t_0) = x_0,\ \forall t \in [t_0-\delta, t_0+\delta],\ |\gamma(t) - x_0| \le \beta \right\}.$$

From the previous calculation, it is clear that $T(Y) \subset Y$. Since Y is closed and convex and $T : Y \to Y$ is continuous (because b is uniformly continuous on X, as X is compact), it suffices to show that T is a compact operator in order to apply the Schauder fixed-point theorem.

Finally, we observe that

$$|T(\gamma)(t_1) - T(\gamma)(t_2)| = \left| \int_{t_0}^{t_1} b(s, \gamma(s))\, ds \right| \le M|t_1 - t_2|.$$

Thus, by the Ascoli-Arzelà theorem, T sends a bounded family of functions to a precompact family, as it is equibounded and equicontinuous.

□

Remark 1.2. As with other proofs in this thesis, the condition that b is continuous on all of $\mathbb{R} \times \mathbb{R}^d$ can be relaxed to continuity on the compact set X; we have presented it in this way for the sake of clarity.

Note that the majority of the results in this section are local, i.e., they depend only on the behaviour of the vector field near the relevant time and space points.

As we will see later, existence is generally easier to prove and holds under mild assumptions; by contrast, the question of uniqueness is more delicate.

We now perform a computation that will have several important consequences. Let $\gamma_1, \gamma_2 \in C^1([t_1, t_2]; \mathbb{R}^d)$ be two solutions of (4), with initial conditions $\gamma_1(t_0) = x_1$ and $\gamma_2(t_0) = x_2$. Suppose there exists $L > 0$ such that for every $t \in [t_1, t_2]$ and for all $x, y \in \mathbb{R}^d$, the vector field satisfies the Lipschitz condition

$$|b(t,x) - b(t,y)| \le L|x-y|.$$

Under this assumption, we can estimate how the difference between the two solutions evolves over time. In particular, we have:

$$\begin{aligned} |\gamma_1(t) - \gamma_2(t)| &= \left| x_1 - x_2 + \int_{t_0}^{t} (b(s,\gamma_1(s)) - b(s,\gamma_2(s)))\, ds \right| \\ &\le |x_1 - x_2| + \int_{t_0}^{t} |b(s,\gamma_1(s)) - b(s,\gamma_2(s))|\, ds \\ &\le |x_1 - x_2| + L \int_{t_0}^{t} |\gamma_1(s) - \gamma_2(s)|\, ds. \end{aligned}$$

Applying Grönwall's lemma, we conclude that

$$|\gamma_1(t) - \gamma_2(t)| \le |x_1 - x_2|\, e^{L|t-t_0|}. \tag{7}$$

In particular, if $x_1 = x_2$, this estimate implies that $\gamma_1(t) = \gamma_2(t)$ for all $t \in [t_1, t_2]$, and thus uniqueness of the solution. This computation proves the so-called *Cauchy–Lipschitz theorem.*

Uniqueness, however, is not guaranteed even if the vector field is continuous and C^∞ except at a single point, as the following classical example shows.

Example 1.3. Let $b(t,x) = \sqrt{|x|}$ then the ODE with initial value $\gamma(0) = 0$ has at least two solutions and both of them are defined over all $\mathbb{R}$. In particular there is the following family of solutions depending on a parameter $c \in [0, +\infty]$:

$$\gamma_c(t) = \begin{cases} 0 & \text{if } t \le c, \\ \frac{1}{4}(t-c)^2 & \text{if } t \ge c. \end{cases}$$

Analogously, for $b(t,x) = |x|^\alpha$ with $0 < \alpha < 1$, uniqueness of solutions fails. Therefore, for $\alpha < 1$, α-Hölder continuity is not sufficient to ensure uniqueness of solutions, whereas for $\alpha = 1$, the Cauchy–Lipschitz theorem guarantees it.

A natural question is whether (local) Lipschitz continuity is necessary for uniqueness of solutions, or if a slightly weaker regularity condition suffices. It turns out that the latter is true: uniqueness holds for functions with an *Osgood modulus of continuity.*

Definition 1.4. A positive (except at zero) function $\omega : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is said to be *Osgood* if it is continuous, increasing, satisfies $\omega(0) = 0$, and the following *Osgood condition* holds:

$$\int_0^1 \frac{ds}{\omega(s)} = +\infty. \tag{8}$$

We define the function $W : (0, 1] \to [0, +\infty)$ as

$$W(x) := \int_x^1 \frac{ds}{\omega(s)}. \tag{9}$$

Notice that W is decreasing, and by the Osgood condition (8) we have

$$\lim_{x\to 0} W(x) = +\infty.$$

Theorem 1.5 Let b be a continuous vector field. Assume that for every open set $\Omega \subset \mathbb{R}^d$, there exist a locally integrable function $\psi(t)$ and an Osgood modulus of continuity ω such that

$$|b(t, x) - b(t, y)| \leq \psi(t)\, \omega(|x - y|) \tag{10}$$

for all $(t, x), (t, y) \in \mathbb{R}\times\Omega$. Then the Cauchy problem associated with b has at most one solution.

Proof. This proof is a generalization of the argument used to prove (7).

Suppose by contradiction that there exist two distinct solutions $\gamma_1, \gamma_2 \in C^1([t_0-\delta, t_0+\delta], \mathbb{R}^d)$ of (5). Define $f(t) := |\gamma_1(t)-\gamma_2(t)|$. Since the solutions are different, there exists a point $z > t_0$ such that $f(z) > 0$. Without loss of generality, we can assume that $f(t) \leq 1$ for all $t \in [t_0, z]$, so that $\omega(f(t)) > 0$ on this interval.

We now compute the derivative of $W(f(t))$. Using the chain rule and the hypothesis (10), we obtain:

$$\frac{d}{dt}W(f(t)) = -\frac{f'(t)}{\omega(f(t))} \geq -\psi(t).$$

Integrating over the interval $[a, z] \subset [t_0, z]$, where $f(a) > 0$, we find:

$$W(f(z)) - W(f(a)) \leq \int_a^z \psi(t)\, dt.$$

Taking the limit as $a \to t_0^+$, we obtain $W(f(z)) - \lim_{a\to t_0^+} W(f(a)) \leq \int_{t_0}^z \psi(t)\, dt < +\infty$. However, by the Osgood condition, $\lim_{x\to 0^+} W(x) = +\infty$, and since $f(t_0) = 0$, we have $\lim_{a\to t_0^+} W(f(a)) = +\infty$, leading to a contradiction.

Therefore, $f(z) = 0$ for all z, and the two solutions must coincide. □

Remark 1.6. A typical example of an Osgood modulus of continuity is

$$\omega(x) := \begin{cases} 0 & \text{if } x = 0, \\ x \log\left(\frac{1}{x}\right) & \text{if } 0 < x < \frac{1}{1000}, \\ \frac{1}{1000}\log(1000) & \text{if } x \geq \frac{1}{1000}. \end{cases}$$

It is easy to check that this function is continuous, increasing, satisfies $\omega(0) = 0$, and fulfills the Osgood condition (8).

Note that what matters in the proof of Theorem 1.5 is the behavior of the function near zero, where the integral $\int_0^1 \frac{ds}{\omega(s)}$ diverges. The values of ω away from zero can be modified freely without affecting the validity of the argument.

Remark 1.7. Notice that when $\omega(x) = x$, we recover the classical Cauchy-Lipschitz theorem in full generality. In other words, if the vector field b belongs to $L^1_{\text{loc}}(\mathbb{R}; W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d))$, then local existence and uniqueness of the solutions hold.

Remark 1.8. From now on, we also impose the condition that ω is concave. This assumption is not very restrictive and is natural for a modulus of continuity. It is not difficult to see that if ω is concave and $\omega(0) = 0$, then, for each $y \leq 1$, we have

$$\omega(x)y \leq \omega(xy).$$

This inequality, which will be used several times in Chapter 4, follows from the fact that the graph of ω at the point xy lies above the line segment connecting the points where the function is evaluated at 0 and x.

2. Flow of the Vector Field

Thanks to the results of the previous section, we know that under suitable assumptions, for each initial point x_0, there exists a unique solution to the ODE (5), at least in a neighborhood of t_0. Moreover, one can show that this unique solution can be extended to a *maximal interval of existence* (a, b), which may be unbounded, i.e., $a = -\infty$ and/or $b = +\infty$.

However, in this thesis, we are not concerned with the technical aspects of extendability or blow-up. Instead, we will focus on vector fields that guarantee the existence of solutions for all times $t \in \mathbb{R}$.

For instance, by examining the proof of Peano's theorem more closely, one can see that if the vector field is continuous and bounded, then for every initial point, it is possible to construct solutions defined on arbitrarily large time intervals. However, since uniqueness is not guaranteed in general, this is not sufficient to ensure the existence of a unique global solution. If, in addition, the vector field satisfies one of the conditions discussed earlier that guarantee uniqueness, such as being in $L^1_{\text{loc}}(\mathbb{R}; W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d))$, then for each initial point, there exists a unique global solution to (5), defined on the entire real line.

Under these assumptions, we define the flow of the vector field.

Definition 1.9. Given a bounded and continuous vector field in $L^1_{\text{loc}}(\mathbb{R}; W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d))$ we define the flow $X(t, x) : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ to be the (unique) function such that

$$\begin{cases} \partial_t X(t, x) = b(t, X(t, x)) \\ X(0, x) = x \end{cases}$$

As a consequence of uniqueness, the following property, known as the *semigroup property of the flow*, holds:

$$X(s, X(t, x)) = X(t + s, x). \tag{11}$$

In particular, if we fix the time and define $X_t := X(t, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$, the semigroup property can be rewritten as

$$X_t \circ X_s = X_{t+s}.$$

Moreover, since $X_0 = \text{Id}$, we see that X_{-t} is the inverse function of X_t. A similar computation to that of (7) shows that in the case where $b \in L^1_{\text{loc}}(\mathbb{R}; W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d))$, the flow is locally Lipschitz continuous in x. In the Osgood case, however, the flow is no longer Lipschitz, but still enjoys a weaker form of continuity.
Moreover, if the vector field b is smooth in the time variable (i.e., C^∞), then the flow is smooth with respect to time. This follows from the observation that in (4), if b is C^1, then the right-hand side is C^1 as a function of time. This implies that the time derivative of the flow is C^1, and therefore the solution is C^2, and so on.

We conclude by stating a theorem which shows that additional spatial regularity of the vector field yields additional spatial regularity of the flow. A proof can be found in Section 1.3 of [**Cri07**].

Theorem 1.10 Let $b : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded and smooth vector field. Then the associated flow $X(t, x)$ is smooth in both time and space. Moreover, for each fixed $t \in \mathbb{R}$, the map

$$X_t : \mathbb{R}^d \to \mathbb{R}^d$$

is a diffeomorphism. More generally, if b is bounded and of class C^1, then the flow is at least of class C^1 in both time and space.

3. Continuity equation in the smooth framework

The continuity equation is a differential equation typically used to describe how a mass distribution evolves over time under a vector field b. In this context, it is natural to consider not only functions as initial data but also measures, which requires giving a distributional meaning to the equation.

Definition 1.11. A family of locally finite signed measures μ_t for $t \in [0,T]$ is said to solve the continuity equation

$$\partial_t \mu_t + \nabla \cdot (b\mu_t) = 0 \tag{12}$$

if, for every test function $\varphi \in C_c^\infty([0,T) \times \mathbb{R}^d)$, the following identity holds:

$$\int_0^T \int_{\mathbb{R}^d} (\partial_t \varphi(t,x) + b(t,x) \cdot \nabla \varphi(t,x)) \, d\mu_t(x) \, dt + \int_{\mathbb{R}^d} \varphi(0,x) \, d\mu_0(x) = 0. \tag{13}$$

Remark 1.12. Note that, in order for (13) to be well-defined, it is sufficient to require

$$\int_0^T \int_K |b(t,x)| \, d|\mu_t|(x) \, dt < +\infty \quad \forall K \subset \mathbb{R}^d, \tag{14}$$

for each compact subset K, without the need for additional regularity assumptions.

Remark 1.13. Under assumption (14), for every test function $\varphi \in C_c^\infty(\mathbb{R}^d)$, the map

$$t \mapsto \int_{\mathbb{R}^d} \varphi(x) \, d\mu_t(x) =: \langle \mu_t, \varphi \rangle$$

is absolutely continuous on $[0,T]$, with derivative

$$\frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) \, d\mu_t(x) = \int_{\mathbb{R}^d} b(t,x) \cdot \nabla \varphi(x) \, d\mu_t(x),$$

which belongs to $L^1([0,T])$. This implies that, for every fixed φ, the map $t \mapsto \langle \mu_t, \varphi \rangle$ admits a unique uniformly continuous representative on $[0,T]$.

Using a density argument, one can then find a unique family $(\tilde{\mu}_t)_{t \in [0,T]}$ of measures such that $\langle \tilde{\mu}_t, \varphi \rangle = \langle \mu_t, \varphi \rangle$ almost everywhere in t and the map $t \mapsto \langle \tilde{\mu}_t, \varphi \rangle$ is uniformly continuous for all $\varphi \in C_c^\infty(\mathbb{R}^d)$. We will always work with this representative, which allows us to evaluate μ_t at every time, including the endpoints $t = 0$ and $t = T$.

Thanks to the previous remarks and definition 13, the Cauchy problem for the continuity equation is well-posed:

$$\begin{cases} \partial_t \mu_t + \nabla \cdot (b\mu_t) = 0 \\ \mu_0 = \bar{\mu}. \end{cases} \tag{15}$$

We will now show that, in the smooth case, there is a unique solution to (15) and provide a precise representation formula for it.

To provide the latter, we first recall the definition of the pushforward of a measure.

Definition 1.14. Given measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, a measurable function $T : X \to Y$, and a measure μ on X, the pushforward of μ, denoted by $T_\# \mu$, is a measure on Y defined as follows:

$$T_\# \mu(B) = \mu(T^{-1}(B)) \quad \text{for all } B \in \mathcal{B}.$$

Using a standard approximation argument, one can show that this condition is equivalent to the following:

$$\int_Y f(y) \, T_\# \mu(dy) = \int_X f(T(x)) \, d\mu(x) \tag{16}$$

for all measurable functions $f : Y \to \mathbb{R}$.

The following proposition provides a solution to (15) and shows that the flow of b transports the initial distribution over time.

Proposition 1.15 If the vector field b belongs to $L^1_{\mathrm{loc}}(\mathbb{R}; W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d))$, is continuous, and bounded, then the family of measures

$$\mu_t = (X_t)_\# \bar{\mu} \tag{17}$$

solves the continuity equation (15).

PROOF. Since the class $C_c^\infty([0,T)) \otimes C_c^\infty(\mathbb{R}^d)$ is dense in $C_c^\infty([0,T)\times\mathbb{R}^d)$ in the C^1 topology, it suffices to check the distributional identity (13) on test functions of the form $\psi(t)\varphi(x)$, with $\psi \in C_c^\infty([0,T))$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$, that is,

$$\int_0^T \psi'(t) \int_{\mathbb{R}^d} \varphi(x)\, d\mu_t(x)\, dt + \int_0^T \psi(t) \int_{\mathbb{R}^d} b(t,x)\cdot\nabla\varphi(x)\, d\mu_t(x)\, dt + \psi(0)\int_{\mathbb{R}^d} \varphi(x)\, d\mu_0(x) = 0.$$

Notice that the map

$$t \mapsto \int_{\mathbb{R}^d} \varphi(x)\, d\mu_t(x) = \int_{\mathbb{R}^d} \varphi(X(t,x))\, d\bar{\mu}(x)$$

is in $C^1([0,T])$, since the flow X is C^1 with respect to the time variable. To check the equation above, we need to show that the distributional derivative of this map is

$$\int_{\mathbb{R}^d} b(t,x)\cdot\nabla\varphi(x)\, d\mu_t(x).$$

Due to the C^1 regularity, we only need to compute the pointwise derivative. Since the flow satisfies

$$\frac{\partial X}{\partial t}(t,x) = b(t, X(t,x)),$$

for every t and x, we can apply the chain rule to obtain

$$\frac{d}{dt}\int_{\mathbb{R}^d} \varphi(X(t,x))\, d\bar{\mu}(x) = \int_{\mathbb{R}^d} \nabla\varphi(X(t,x))\cdot b(t,X(t,x))\, d\bar{\mu}(x) = \int_{\mathbb{R}^d} b(t,x)\cdot\nabla\varphi(x)\, d\mu_t(x).$$

Hence, we have shown the desired result. □

Remark 1.16. If $\bar{\mu} = \bar{\rho}\mathscr{L}^d$, where $\bar{\rho} \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, and $\mu_t = (X_t)_\#\bar{\mu}$, then μ_t is absolutely continuous with respect to the Lebesgue measure, and we can write

$$\mu_t = \rho_t\, \mathscr{L}^d,$$

where the density ρ_t is given by the formula

$$\rho_t(x) = \bar{\rho}(X(-t,x)) \cdot \left|\det \nabla X(-t,x)\right|. \tag{18}$$

A detailed proof can be found in Section 2 of [**AC14**]; the formula is derived using the push-forward definition and the area formula.

Finally we prove that (17) is the unique solution when the vector field is smooth enough.

Proposition 1.17 The representation formula (17) gives the unique solution of (15) when the vector field is bounded and C^1.

PROOF. Let μ_t be a solution to the continuity equation. We will show that $\mu_t = (X_t)_\#\bar{\mu}$. Notice that $\mu_t = (X_t)_\#\bar{\mu}$ if and only if $(X_{-t})_\#\mu_t = \bar{\mu}$. Thus, it is natural to define $\tilde{\mu}_t := (X_{-t})_\#\mu_t$ and prove that it coincides with $\bar{\mu}$. Since $\tilde{\mu}_0 = \bar{\mu}$, it suffices to show that the distributional derivative is zero, i.e.,

$$\int_0^T \psi'(t)\int_{\mathbb{R}^d} \varphi(x)\, d\tilde{\mu}_t(x)\, dt = -\int_{\mathbb{R}^d} \psi(0)\varphi(x) d\bar{\mu}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $\psi \in C_c^\infty([0,T))$. Rewriting the inner integral using the definition of pushforward, we have

$$\int_{\mathbb{R}^d} \varphi(x)\, d\tilde{\mu}_t(x) = \int_{\mathbb{R}^d} \varphi(X(-t,x))\, d\mu_t(x).$$

To proceed, we apply the hypothesis that μ_t is a solution to the continuity equation with the test function $P(t,x) = \psi(t)\varphi(X(-t,x))$, which gives

$$\int_0^T \int_{\mathbb{R}^d} \left(\partial_t P(t,x) + b(t,x)\cdot \nabla P(t,x)\right) d\mu_t(x)\, dt = -\int_{\mathbb{R}^d} \psi(0)\varphi(x) d\bar{\mu}. \tag{19}$$

Note that

$$\partial_t P(t,x) = \psi'(t)\varphi(X(-t,x)) - \psi(t)\nabla\varphi(X(-t,x))\cdot b(-t, X(-t,x))$$

and

$$\nabla P(t,x) = \psi(t)\nabla\varphi(X(-t,x))\cdot \nabla X(-t,x).$$

Substituting these expressions into (19), we get that the thesis is equivalent to proving

$$\int_{\mathbb{R}^d} -\nabla\varphi(X(-t,x))\cdot b(-t,X(-t,x)) + b(t,x)\nabla\varphi(X(-t,x))\cdot\nabla X(-t,x)\, d\mu_t(x) = 0.$$

To conclude the proof, it suffices to show that

$$b(t,x)\nabla X(-t,x) = b(-t, X(-t,x)). \tag{20}$$

This follows by differentiating the semigroup law

$$X(t, X(-t,x)) = x$$

with respect to t. Using the chain rule, we have

$$\frac{d}{dt}X(t,X(-t,x)) = \partial_t X(t,X(-t,x)) + \nabla X(t,X(-t,x))\frac{d}{dt}X(-t,x) = 0.$$

By the properties of the flow (i.e., the semigroup law), we can rewrite this expression as

$$b(t,x) = \nabla X(t,X(-t,x))b(-t,X(-t,x)).$$

Finally, we conclude by noting that

$$\nabla X(t,X(-t,x)) = \left(\nabla X(-t,x)\right)^{-1}$$

since $X(t,\cdot)$ and $X(-t,\cdot)$ are inverses. □

4. Transport Equation in the Smooth Framework

In contrast to the continuity equation, the transport equation transports a scalar quantity through the flow. Thus, we consider functions rather than measures. In the smooth case, we will deal with the Cauchy problem in the classical sense.

$$\begin{cases} \partial_t u + b\cdot\nabla u = 0 \\ u(0,x) = \bar{u}(x) \end{cases} \qquad u : [0,T]\times\mathbb{R}^d \to \mathbb{R}, \tag{21}$$

A crucial observation is as follows: let u be a solution of (21), and consider the composition $u(t, X(t,x))$. Differentiating it with respect to time, we obtain

$$\frac{d}{dt}u(t,X(t,x)) = \partial_t u(t,X(t,x)) + b(t,X(t,x))\cdot\nabla u(t,X(t,x)) = 0,$$

where we applied the chain rule and used the fact that u is a solution of (21). This computation, typical of the theory of characteristics, shows that u is constant along the characteristic curves of the vector field. In particular, we have

$$u(t,X(t,x)) = u(0,X(0,x)) = u(0,x) = \bar{u}(x),$$

which implies, using the inverse flow,

$$u(t,x) = \bar{u}(X(-t,x)). \tag{22}$$

This equation shows that, under the hypothesis that $\bar{u}\in C^1$ and the vector field b is bounded, continuous, and locally Lipschitz, there is at most a unique solution to (21), and it is given by (22). We now prove that this is indeed a solution to (21). First, by construction, we have $u(0,x) = \bar{u}(x)$. For the other equation, we compute

$$\partial_t u + b\cdot\nabla u = \nabla\bar{u}(X(-t,x))\cdot(-b(t,X(-t,x))) + b(t,x)\nabla\bar{u}(X(-t,x))\nabla X(-t,x) = 0,$$

where the last equality follows from (20), which was proven in the previous proof.

We have thus proved the following proposition.

Proposition 1.18 (Existence, Uniqueness, and Representation Formula for the Transport Equation) Let $\bar{u} \in C^1(\mathbb{R}^d)$, and let $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded, C^1 vector field. Then the Cauchy problem (21) admits the following properties:

(1) There exists a unique classical solution $u \in C^1([0,T] \times \mathbb{R}^d)$.

(2) The solution is given by the representation formula

$$u(t,x) = \bar{u}(X(-t,x)),$$

where $X(t,x)$ denotes the flow of the vector field b, i.e., the unique solution to

$$\frac{d}{dt}X(t,x) = b(t, X(t,x)), \quad X(0,x) = x.$$

(3) The solution u is constant along the flow lines of b, in the sense that

$$\frac{d}{dt}u(t, X(t,x)) = 0.$$

5. Divergence-Free Vector Fields

In the last section of the chapter, we add the assumption that the vector field has zero divergence (with respect to the spatial variables). This condition will lead to several important consequences and will be assumed throughout the remainder of the thesis.

First, observe that for divergence-free vector fields, and assuming sufficient regularity of both the vector field b and the function f, we have

$$\nabla \cdot (bf) = (\nabla \cdot b)f + b \cdot \nabla f = b \cdot \nabla f.$$

As a result, the continuity equation (2) and the transport equation (3) become the same equation. In this section, we will use the representation formulas to prove that the vector field is incompressible.

Proposition 1.19 If b is a divergence-free vector field, bounded and C^1, then the flow preserves the measure. That is, for all $t \in [0,T]$, we have

$$(X_t)_\# \mathscr{L}^d = \mathscr{L}^d.$$

Proof. We compute μ_t, the solution of the equation, in two different ways, starting from $\mu_0 = dx$, where dx is the Lebesgue measure. By adapting the representation formula for the transport equation (22) to this case, we deduce that $\mu_t = dx$ for all $t \in [0,T]$.

On the other hand, from (18), we know that

$$1 = |\det \nabla X(-t,x)|\,.$$

Therefore, for all $t \in [0,T]$ and $x \in \mathbb{R}^d$, we have $|\det \nabla X(t,x)| = 1$.

By changing variables in the Lebesgue integral, we have, for every measurable function f,

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_{\mathbb{R}^d} f(X(t,x))|\det \nabla X(t,x)|\,dx = \int_{\mathbb{R}^d} f(X(t,x))\,dx.$$

By the definition of pushforward, this last equation gives

$$(X_t)_\# \mathscr{L}^d = \mathscr{L}^d.$$

as desired. □

CHAPTER 2

DiPerna–Lions Classical Theory in the Sobolev Case

1. Introduction

In the first part of this chapter, we introduce the framework in which we revisit the results and questions from Chapter 1, now under weaker regularity assumptions on the vector field. The central questions are whether one can still establish existence and uniqueness of solutions, and in what sense these notions should be understood.

A major breakthrough on these questions was provided by DiPerna and Lions in their seminal paper [**DL89**], where they developed a well-posedness theory for the transport and continuity equations associated with Sobolev vector fields. These PDEs also played a crucial role in establishing existence and uniqueness results for ordinary differential equations driven by non-smooth vector fields.

From the previous chapter, two important observations can be made. First, as illustrated by Example 1.3, uniqueness is lost when the regularity of the vector field drops below the Osgood/Lipschitz threshold. Therefore, in order to construct a meaningful analogue of the classical flow of a vector field, one must go beyond simply verifying that it solves the equation in a weak sense. Second, this issue is also closely tied to the continuity equation, since there is still hope that a representation formula similar to (17) might hold, even in settings of lower regularity.

We will focus on bounded, divergence-free vector fields in $L^1_{\mathrm{loc}}(\mathbb{R}, W^{1,p}_{\mathrm{loc}}(\mathbb{R}^d, \mathbb{R}^d))$, where $p \geq 1$ is fixed. Moreover, we will primarily consider the Cauchy problem with initial data given by functions (rather than measures), interpreting the equation in the distributional sense. The following definition is (13) specialized to functions (i.e take $\mu_t = u(t,\cdot)\mathscr{L}^d$). From a conceptual point of view, even though in the divergence-free case the continuity equation coincides with the transport equation, we refer to it as the *continuity equation* when the initial datum is a measure, whereas the term *transport equation* is more appropriate when the initial datum is a scalar function.

Definition 2.1. We say that a locally summable function $u(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is a solution of the Cauchy problem for the transport equation with initial datum $\bar{u} \in L^\infty_{\mathrm{loc}}(\mathbb{R}^d)$ if it satisfies the system in the distributional sense:

$$\begin{cases} \partial_t u + \nabla \cdot (bu) = 0, \\ u(0,\cdot) = \bar{u}, \end{cases} \tag{23}$$

that is, if for every test function $\varphi \in C^\infty_c([0,T) \times \mathbb{R}^d)$, the following identity holds:

$$\int_0^T \int_{\mathbb{R}^d} u\,[\partial_t \varphi + b \cdot \nabla \varphi]\; dx\,dt = -\int_{\mathbb{R}^d} \bar{u}(x)\varphi(0,x)\,dx. \tag{24}$$

Remark 2.2. In analogy with Remark 1.13, any weak solution $u \in L^\infty([0,T] \times \mathbb{R}^d)$ admits a representative that is weakly-* continuous in time with values in $L^\infty(\mathbb{R}^d)$. More precisely, for every $\varphi \in L^1(\mathbb{R}^d)$, the map

$$t \mapsto \int_{\mathbb{R}^d} u(t,x)\varphi(x)\,dx$$

is continuous on $[0,T]$, which uniquely identifies the values of $u(t,\cdot)$ in the weak-* topology for all $t \in [0,T]$.

For the ODE, on the other hand, one would be satisfied if it can be shown that there exists a substitute for the flow, as follows.

Definition 2.3. A *regular Lagrangian flow* is a map $X(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying the following conditions:

1) For almost every x, the map $t \mapsto X(t,x)$ is absolutely continuous;
2) $X(0,x) = x$, and for almost every x, the map $t \mapsto X(t,x)$ satisfies $\frac{d}{dt}X(t,x) = b(t, X(t,x))$ for almost every $t \in [0,T]$;
3) For every $t \in [0,T]$, we have $X(t,\cdot)_{\#}\mathscr{L}^d = \mathscr{L}^d$.

Remark 2.4. This definition is particularly natural in the divergence-free case, as discussed in Section 5 of Chapter 1. In the classical Lipschitz setting, the flow indeed satisfies the third property as shown in Proposition 1.19. Requiring preservation of the Lebesgue measure is crucial, as it ensures that trajectories do not concentrate or collapse onto sets of measure zero. This phenomenon may occur, for instance, when $b(t,x) = \sqrt{|x|}$, where the flow could collapse mass at the origin. The other conditions are appropriate for the low-regularity regime, where one cannot expect the ODE to hold pointwise everywhere.

Remark 2.5. Notice that the third condition is also important because it ensures that the notion of solution is invariant under modifications of b on sets of negligible measure. Specifically, if

$$b(t,x) = \tilde{b}(t,x) \quad \text{for} \quad \mathscr{L}^{d+1}\text{-a.e.}\,(t,x) \in [0,T] \times \mathbb{R}^d,$$

then it is straightforward to verify that X is a regular Lagrangian flow with respect to b if and only if it is a regular Lagrangian flow with respect to $\tilde{b}$.

Remark 2.6. In the general case of vector fields with non-zero divergence, the third condition is typically weakened: we require the existence of a constant $L > 0$ such that

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

in the sense of measures, for every $t \in [0,T]$. This accounts for the possible compression or expansion induced by the divergence of the field.

In Section 2, we will show how DiPerna and Lions proved uniqueness for the PDE. In Section 3, we will explain how their results can be used to deduce the existence and uniqueness of a regular Lagrangian flow.

2. Renormalized solutions

The main idea for dealing with the PDE in the low-regularity setting is to regularize both the vector field and the solutions, and then examine how the properties they possess when smoothed can be passed to the limit.
First, we will show that the existence of solutions to (24) holds under mild assumptions on the vector field and the initial datum.

Proposition 2.7 Let b be a bounded, divergence-free vector field in $L^1_{\text{loc}}([0,T], L^1_{\text{loc}}(\mathbb{R}^d, \mathbb{R}^d))$, and let $\bar{u} \in L^\infty$. Then, there exists a solution to (24).

Proof. We begin by mollifying both the vector field and the initial datum. Let ρ_ε and χ_ε be two families of mollifiers on $\mathbb{R}^d$ and $\mathbb{R} \times \mathbb{R}^d$, respectively. We define the regularizations $\bar{u}_\varepsilon = \bar{u} * \rho_\varepsilon$ and $b_\varepsilon = b * \chi_\varepsilon$. Consider the equation with the regularized vector field and initial datum, and let u_ε be its unique smooth solution:

$$\begin{cases} \partial_t u_\varepsilon + \nabla \cdot (b_\varepsilon u_\varepsilon) = 0, \\ u_\varepsilon(0,\cdot) = \bar{u}_\varepsilon. \end{cases}$$

Two observations are crucial. First, since $\bar{u} \in L^\infty$, the mollified initial datum $\bar{u}_\varepsilon$ is uniformly bounded. Second, in the smooth setting, solutions to the transport equation are uniformly bounded by the L^∞-norm of the initial datum (see, for instance, the representation formula (22)). Therefore, $\|u_\varepsilon\|_{L^\infty}$ remains uniformly bounded. It follows that there exists a subsequence (not relabeled) that converges weak-* in $L^\infty([0,T]\times\mathbb{R}^d)$ to some function u. Now, fix a test function $\varphi\in C_c^\infty([0,T)\times\mathbb{R}^d)$. Since the u_ε are smooth solutions, integration by parts gives

$$\int_0^T\int_{\mathbb{R}^d} u_\varepsilon\left(\partial_t\varphi+b_\varepsilon\cdot\nabla\varphi\right)dx\,dt=-\int_{\mathbb{R}^d}\bar u_\varepsilon(x)\varphi(0,x)\,dx.$$

Finally, using the strong convergence of $\bar u_\varepsilon\to\bar u$ and $b_\varepsilon\to b$, together with the weak-* convergence of $u_\varepsilon\to u$, we can pass to the limit in the above identity and obtain

$$\int_0^T\int_{\mathbb{R}^d} u\left(\partial_t\varphi+b\cdot\nabla\varphi\right)dx\,dt=-\int_{\mathbb{R}^d}\bar u(x)\varphi(0,x)\,dx,$$

which shows that u is a distributional solution. □

In the direction of proving uniqueness, a crucial definition arises from the following observation. By (22), the unique solution of the transport equation in the smooth setting is given by

$$u(t,x)=\bar u(X(-t,x)).$$

Notice that if we have a C^1 function f, and if one is interested in finding the solution $v(t,x)$ starting from $f(\bar u)$, then, by the representation formula, it is simply

$$v(t,x)=f(u(t,x)).$$

Mimicking this simple observation in the non-smooth setting will be crucial for proving uniqueness.

Definition 2.8. Let b be a bounded, divergence-free vector field in $L^1_{\mathrm{loc}}([0,T],L^1_{\mathrm{loc}}(\mathbb{R}^d,\mathbb{R}^d))$. We say that a bounded distributional solution u of the transport equation is renormalized if, for every $\beta\in C^1(\mathbb{R})$, $\beta(u)$ is a distributional solution of

$$\begin{cases}\partial_t(\beta(u))+\nabla\cdot(b\beta(u))=0,\\ u(0,\cdot)=\beta(\bar u).\end{cases}$$

Furthermore, if every bounded distributional solution is renormalized, the vector field b is said to have the renormalization property.

If a vector field has the renormalization property, under mild assumptions, it suffices to prove uniqueness in a certain class of functions. Heuristically, one would like to prove an identity of the form

$$\int_{\mathbb{R}^d}u(t,x)^2\,dx=0,$$

which would immediately imply $u=0$. This is motivated by the formal computation

$$\frac{d}{dt}\int_{\mathbb{R}^d}u(t,x)^2\,dx=2\int_{\mathbb{R}^d}\partial_t u(t,x)\,u(t,x)\,dx=-2\int_{\mathbb{R}^d}u(t,x)\,b(t,x)\cdot\nabla u(t,x)\,dx=0,$$

where the last equality follows from the fact that the vector field b is divergence-free, and from the identity

$$-2u(t,x)\,\nabla u(t,x)=-\nabla\big(u(t,x)^2\big).$$

The proof consists of making this computation rigorous within the distributional framework. We will use the following distributional version of Gronwall's lemma.

Lemma 2.9 Let $f\in C^0([0,T])$ be a continuous function such that $f(0)=0$, and assume that for every non-negative test function $\psi\in C_c^\infty((0,T))$ the following inequality holds:

$$\int_0^T f(s)\psi'(s)\,ds\ge 0.$$

Then $f(t)\le 0$ for all $t\in[0,T]$.

Proof. Let $x, y \in (0,T)$ with $x < y$. We will show that $f(x) \geq f(y)$, and since $f(0) = 0$, this will imply that $f(t) \leq 0$ for all $t \in [0,T]$.

Let $\varepsilon > 0$ be small enough such that $x - \varepsilon > 0$ and $y + \varepsilon < T$. Define a Lipschitz function ψ supported in $(0,T)$ as follows:

$$\psi(t) = \begin{cases} 0 & \text{if } t \leq x - \varepsilon, \\ \frac{t-(x-\varepsilon)}{\varepsilon} & \text{if } t \in (x-\varepsilon, x), \\ 1 & \text{if } t \in (x, y), \\ \frac{(y+\varepsilon)-t}{\varepsilon} & \text{if } t \in (y, y+\varepsilon), \\ 0 & \text{if } t \geq y + \varepsilon. \end{cases}$$

This function can be approximated in $C_c^\infty((0,T))$ by smooth non-negative functions, so the inequality

$$\int_0^T f(s)\psi'(s)\, ds \geq 0$$

also holds for ψ. Computing the integral, we get

$$\int_0^T f(s)\psi'(s)\, ds = \frac{1}{\varepsilon}\int_{x-\varepsilon}^{x} f(s)\, ds - \frac{1}{\varepsilon}\int_{y}^{y+\varepsilon} f(s)\, ds.$$

Hence,

$$\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x} f(s)\, ds \geq \frac{1}{\varepsilon}\int_{y}^{y+\varepsilon} f(s)\, ds.$$

Since f is continuous, we may let $\varepsilon \to 0$ and obtain

$$f(x) \geq f(y).$$

Since this holds for all $x < y$ in $(0,T)$, and $f(0) = 0$, we conclude that $f(t) \leq 0$ for all $t \in [0,T]$. □

Proposition 2.10 If b is bounded and divergence-free, and it has the renormalization property, then uniqueness holds in the class of bounded solutions.

Proof. Suppose we have two bounded solutions v and w corresponding to the same initial datum. Since the equation is linear, we can consider their difference $u = v - w$, which is also a bounded solution with zero initial datum. Therefore, it suffices to prove uniqueness in the case where the initial datum is zero; that is, any bounded solution with zero initial datum must be identically zero. Let u be such a solution, and consider the function $\beta(x) = x^2$. By the renormalization property, it follows that for every test function $\varphi_1 \in C_c^\infty([0,T) \times \mathbb{R}^d)$, we have

$$\int_0^T \int_{\mathbb{R}^d} u^2 \left(\partial_t \varphi_1 + b \cdot \nabla \varphi_1\right)\, dx\, dt = 0. \tag{25}$$

Now, fix $\psi(t) \in C_c^\infty((0,T))$, $\varphi(t,x) \in C_c^\infty([0,T) \times \mathbb{R}^d)$ such that $\psi(t) \geq 0$ and

$$\partial_t \varphi(t,x) \leq -\|b\|_\infty |\nabla \varphi(t,x)| \quad \text{on } [0,T] \times \mathbb{R}^d.$$

Test (25) with $\psi(t)\varphi(t,x)$, and let $f(t) := \int_{\mathbb{R}^d} u(t,x)^2 \varphi(t,x)\, dx$. The equation can be rewritten as

$$-\int_0^T f(t)\psi'(t)\, dt = \int_0^T \int_{\mathbb{R}^d} u(t,x)^2 \psi(t) \left(\partial_t \varphi(t,x) + b(t,x) \cdot \nabla \varphi(t,x)\right)\, dx\, dt.$$

By the choice of φ and ψ, the second term is non-negative, so we obtain

$$\int_0^T f(t)\psi'(t)\, dt \geq 0.$$

Notice that f is continuous since u^2 is also a solution of the transport equation, and every solution in L^∞ is weak-* continuous in time. In particular, for any fixed test function $\varphi(t,x)$, the map $t \mapsto f(t) = \int_{\mathbb{R}^d} u(t,x)^2 \varphi(t,x)\, dx$ is continuous. Moreover, $f(0) = 0$, so by Lemma 2.9, we conclude that $f(t) \leq 0$ for all $t \in [0,T]$.

Suppose, in addition, that φ satisfies the following properties for some fixed $t_1 \in (0,T)$ and $R > 0$:

- $\varphi(t_1, x) \geq 0$ for all $x \in \mathbb{R}^d$,
- $\varphi(t_1, x) = 1$ on $B_R(0)$, the ball of radius R centered at the origin.

Given these assumptions, we can proceed as follows: Since $f(t_1) \leq 0$, we obtain:

$$\int_{B_R(0)} u(t_1, x)^2 \, dx \leq 0,$$

and by the arbitrariness of t_1 and R, this implies that $u = 0$ almost everywhere, as desired. We will now proceed to construct the desired function φ in the next lemma. □

With the proof of the following lemma, the proof of Proposition 2.10 is complete.

Lemma 2.11 Fix $t_1 \in (0,T)$ and $R > 0$. Then, there exists $\varphi \in C_c^\infty([0,T) \times \mathbb{R}^d)$ such that:

- $\partial_t \varphi(t,x) \leq -\|b\|_\infty |\nabla \varphi(t,x)|$ on $[0,T] \times \mathbb{R}^d$,
- $\varphi(t_1, x) \geq 0$ for all $x \in \mathbb{R}^d$,
- $\varphi(t_1, x) = 1$ on $B_R(0)$, the ball of radius R centered at the origin.

Proof. Let $\delta > 0$ be a small parameter and define the smooth function

$$\rho(x) = \sqrt{|x|^2 + \delta^2},$$

which is smooth on $\mathbb{R}^d$ and closely approximates $|x|$ for large $|x|$.

Next, let $\psi \in C^\infty(\mathbb{R})$ be a smooth function that satisfies:

- $\psi(s) = 1$ for $s \leq R$,
- $\psi(s) = 0$ for $s \geq R + 1$,
- $\psi'(s) \leq 0$ for all $s \in \mathbb{R}$ (i.e. a decreasing function).

We now define the function $\varphi(t,x)$ as follows:

$$\varphi(t,x) = \psi\left(\rho(x) - \|b\|_\infty K(t_1 - t)\right),$$

for $(t,x) \in [0,T) \times \mathbb{R}^d$, where K is a large constant to be chosen later.

Since ρ is smooth, clearly φ is smooth. In addition, by choosing K large enough (it suffices that $\|b\|_\infty K(T - t_1) > R + 1$), the function is also compactly supported.

Next, we verify that $\varphi(t,x)$ satisfies the three required conditions. We need to show that

$$\partial_t \varphi(t,x) \leq -\|b\|_\infty |\nabla \varphi(t,x)|.$$

First, compute the time derivative of $\varphi(t,x)$:

$$\partial_t \varphi(t,x) = K\|b\|_\infty \psi'\left(\rho(x) - \|b\|_\infty K(t_1 - t)\right).$$

Next, we compute the spatial gradient of $\varphi(t,x)$ as follows:

$$\nabla \varphi(t,x) = \psi'\left(\rho(x) - K\|b\|_\infty (t_1 - t)\right) \nabla \rho(x).$$

Thus, we have

$$|\nabla \varphi(t,x)| = \left|\psi'\left(\rho(x) - K\|b\|_\infty (t_1 - t)\right)\right| |\nabla \rho(x)| \leq \left|\psi'\left(\rho(x) - K\|b\|_\infty (t_1 - t)\right)\right|.$$

Finally, since $\psi'(s) \leq 0$, we have

$$\psi'\left(\rho(x) - K\|b\|_\infty (t_1 - t)\right) = -\left|\psi'\left(\rho(x) - K\|b\|_\infty (t_1 - t)\right)\right|.$$

Thus, we obtain

$$\partial_t \varphi(t,x) = -K\|b\|_\infty \left|\psi'\left(\rho(x) - K\|b\|_\infty (t_1 - t)\right)\right| = -K\|b\|_\infty |\nabla \varphi(t,x)|,$$

which gives us the desired inequality.

Finally, we prove the last two conditions. At $t = t_1$, we have

$$\varphi(t_1, x) = \psi\left(\rho(x) - K\|b\|_\infty (t_1 - t_1)\right) = \psi(\rho(x)).$$

Since $\psi(s) \geq 0$ for all s, we conclude that

$$\varphi(t_1, x) \geq 0 \quad \text{for all } x \in \mathbb{R}^d.$$

For the last condition, notice that for $x \in B_R(0)$, we have $|x| \leq R$, and hence

$$\rho(x) = \sqrt{|x|^2 + \delta^2} \leq \sqrt{R^2 + \delta^2}.$$

By choosing $\delta > 0$ sufficiently small, we can ensure that $\rho(x) \leq R + 1$ for all $x \in B_R(0)$, and thus

$$\varphi(t_1, x) = \psi(\rho(x)) = 1 \quad \text{for all } x \in B_R(0),$$

as wanted.

□

In [**DL89**], DiPerna and Lions introduced the notion of renormalized solutions and proved the following crucial theorem:

Theorem 2.12 (DiPerna–Lions) If b is a vector field that belongs to $L^1_{\mathrm{loc}}(\mathbb{R}, W^{1,p}_{\mathrm{loc}}(\mathbb{R}^d, \mathbb{R}^d))$, is bounded, and is divergence-free, then it has the renormalization property.

We will prove this theorem by combining several key steps.

The main idea is to regularize the solution and analyze the error term that arises when the regularized function is treated as a solution to the equation. Let u be a bounded solution of the transport equation, and let ρ_ε be a mollifier on $\mathbb{R}^d$ with ρ supported on the unit ball and radial. Throughout this section, all convolutions are taken with respect to the spatial variables only. We denote the mollified function $u * \rho_\varepsilon$ by u_ε. We define the error term $r_\varepsilon(t, x)$ as the function satisfying

$$\partial_t u_\varepsilon + b \cdot \nabla u_\varepsilon = r_\varepsilon, \tag{26}$$

in the sense of distributions. Our goal now is to better understand its structure.

The following proposition is motivated by this formal computation:

$$\partial_t u_\varepsilon + b \cdot \nabla u_\varepsilon = \partial_t(u * \rho_\varepsilon) + b \cdot \nabla u_\varepsilon = (-b \cdot \nabla u) * \rho_\varepsilon + b \cdot \nabla u_\varepsilon = b \cdot \nabla u_\varepsilon - (b \cdot \nabla u)_\varepsilon.$$

Proposition 2.13 In the sense of distributions, in the divergence free vector field case, we have the identity

$$r_\varepsilon = b \cdot \nabla u_\varepsilon - (b \cdot \nabla u)_\varepsilon. \tag{27}$$

We refer to r_ε as the *(DiPerna–Lions) commutator*, as it measures the discrepancy between applying the gradient before or after convolution. Furthermore, the commutator can be rewritten as

$$r_\varepsilon(t, x) = \int_{\mathbb{R}^d} u(t, x + \varepsilon h) \, \frac{b(t, x + \varepsilon h) - b(t, x)}{\varepsilon} \cdot \nabla \rho(h) \, dh, \tag{28}$$

where the equality holds pointwise almost everywhere.

Proof. Let, as usual, $\varphi \in C^\infty_c([0, T) \times \mathbb{R}^d)$ be a test function. By definition, we have

$$\int_0^T \int_{\mathbb{R}^d} u_\varepsilon \, \partial_t \varphi + u_\varepsilon \, b \cdot \nabla \varphi \, dx \, dt + \int_{\mathbb{R}^d} u_\varepsilon(0, x) \, \varphi(0, x) \, dx = -\int_0^T \int_{\mathbb{R}^d} r_\varepsilon \, \varphi \, dx \, dt. \tag{29}$$

On the other hand, since u is a solution, we have a similar identity. We test it against $\varphi * \rho_\varepsilon$, which is an admissible test function because ρ has compact support, and therefore $\varphi * \rho_\varepsilon$ remains compactly supported. We then obtain:

$$\int_0^T \int_{\mathbb{R}^d} u \, \partial_t(\varphi * \rho_\varepsilon) + u \, b \cdot \nabla(\varphi * \rho_\varepsilon) \, dx \, dt + \int_{\mathbb{R}^d} u(0, x) \, (\varphi * \rho_\varepsilon)(0, x) \, dx = 0. \tag{30}$$

We want to rewrite the second equation and substitute it into the first. First, by exploiting the definition of convolution, Fubini's theorem, and the fact that ρ is radial, we compute:

$$\begin{aligned}\int_{\mathbb{R}^d} u\,\partial_t(\varphi * \rho_\varepsilon)\,dx &= \int_{\mathbb{R}^d} u\,(\partial_t\varphi * \rho_\varepsilon)\,dx \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} u(t,x)\,\partial_t\varphi(t,x-y)\,\rho_\varepsilon(y)\,dy\,dx \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} u(t,x)\,\partial_t\varphi(t,x-y)\,\rho_\varepsilon(-y)\,dy\,dx.\end{aligned}$$

Now we perform the change of variables $x = z + y$, so that $dx = dz$, and obtain:

$$\begin{aligned}\int_{\mathbb{R}^d} u\,\partial_t(\varphi * \rho_\varepsilon)\,dx &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} u(t,z+y)\,\partial_t\varphi(t,z)\,\rho_\varepsilon(-y)\,dz\,dy \\ &= \int_{\mathbb{R}^d} u_\varepsilon(t,z)\,\partial_t\varphi(t,z)\,dz.\end{aligned}$$

We can now use this last equality to combine (29) and (30), yielding:

$$\begin{aligned}-\int_0^T\int_{\mathbb{R}^d} r_\varepsilon\,\varphi\,dx\,dt &= \int_0^T\int_{\mathbb{R}^d} u_\varepsilon\, b\cdot\nabla\varphi - u\, b\cdot\nabla(\varphi * \rho_\varepsilon)\,dx\,dt \\ &\quad + \int_{\mathbb{R}^d} u_\varepsilon(0,x)\,\varphi(0,x) - u(0,x)\,(\varphi * \rho_\varepsilon)(0,x)\,dx.\end{aligned}$$

Finally, observe that

$$\int_{\mathbb{R}^d} u_\varepsilon(0,x)\,\varphi(0,x)\,dx = \int_{\mathbb{R}^d} u(0,x)\,(\varphi * \rho_\varepsilon)(0,x)\,dx$$

by the same computation as above. We now want to make use of the fact that u_ε is smooth and that the vector field b is divergence-free. By integration by parts, we have:

$$\int_{\mathbb{R}^d} u_\varepsilon\, b\cdot\nabla\varphi\,dx = -\int_{\mathbb{R}^d} \varphi\, b\cdot\nabla u_\varepsilon\,dx, \tag{31}$$

where we used that $\nabla\cdot b = 0$ and that φ has compact support.

Since we are in the divergence-free case, it is important to underline that distributionally, $(b\cdot\nabla u)_\varepsilon$ tested by φ is equal to:

$$\int_0^T\int_{\mathbb{R}^d}((bu) * \rho_\varepsilon)\cdot\nabla\varphi\,dx\,dt,$$

as can be seen by formally applying integration by parts and observing how differentiation and convolution interact. Therefore, to prove (27), it suffices to show that

$$\int_0^T\int_{\mathbb{R}^d}((bu) * \rho_\varepsilon)\cdot\nabla\varphi\,dx\,dt = \int_0^T\int_{\mathbb{R}^d} u\, b\cdot\nabla(\varphi * \rho_\varepsilon)\,dx\,dt,$$

which can be proved in the same way as an analogous equality was proved for the time derivative. Finally, we show (28). Until now, we have shown that

$$\int_0^T\int_{\mathbb{R}^d} r_\varepsilon\,\varphi\,dx\,dt = \int_0^T\int_{\mathbb{R}^d} \varphi\, b\cdot\nabla u_\varepsilon\,dx\,dt + \int_0^T\int_{\mathbb{R}^d}((bu) * \rho_\varepsilon)\cdot\nabla\varphi\,dx\,dt.$$

We want to obtain cancellations on the right-hand side, and the next steps will be motivated by this goal. Using (31), we have

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^d}(((bu) * \rho_\varepsilon) - b(u * \rho_\varepsilon))\cdot\nabla\varphi\,dx\,dt \\ &= \int_0^T\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \nabla\varphi(t,x)\,\rho_\varepsilon(x-y)\,u(t,y)\,(b(t,x) - b(t,y))\,dx\,dy\,dt \\ &= \int_0^T\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi(t,x)\,\nabla\rho_\varepsilon(x-y)\,u(t,y)\,(b(t,x) - b(t,y))\,dx\,dy\,dt.\end{aligned}$$

where in the last line we used that b is divergence-free and applied integration by parts. By the arbitrariness of φ, we can conclude, by the fundamental lemma of the calculus of variations, that for almost every (t,x) we have:

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

Finally, by changing variables according to $h = \frac{y-x}{\varepsilon}$, and using the fact that ρ is radial, we obtain (28). □

Before proving Theorem 2.12, we need one last crucial proposition: namely, that as $\varepsilon \to 0$, the commutator vanishes in L^1_{loc}.

Proposition 2.14 If b is a vector field that belongs to $L^1_{\text{loc}}(\mathbb{R}, W^{1,p}_{\text{loc}}(\mathbb{R}^d,\mathbb{R}^d))$, is bounded, and it is divergence-free, then

$$r_\varepsilon \longrightarrow 0 \quad \text{in } L^1_{\text{loc}}([0,T)\times\mathbb{R}^d) \quad \text{as } \varepsilon \to 0.$$

Proof. By applying (28), the dominated convergence theorem, and the fact that

$$\frac{b(t,x+\varepsilon h) - b(t,x)}{\varepsilon} \longrightarrow \nabla b(t,x)h \quad \text{in } L^1_{\text{loc}} \text{ as } \varepsilon \to 0 \tag{32}$$

for a Sobolev vector field, it suffices to show that

$$\int_{\mathbb{R}^d} u(t,x)\, (\nabla b(t,x)h)\cdot\nabla\rho(h)\, dh = 0.$$

Finally, by noticing that, through integration by parts, we have

$$\int_{\mathbb{R}^d} z_i \frac{\partial\rho(z)}{\partial z_j}\, dz = -\delta_{ij},$$

where δ_{ij} is the Kronecker delta, the desired integral is zero because the divergence of b is zero. □

We finally prove Theorem 2.12. One can see that the crucial fact is the convergence of the commutator to zero in the strong sense (L^1), and it is the only part of the the proof where Sobolev regularity is used. If, in another context, one is able to prove that the commutator goes to zero, then this scheme allows one to prove that bounded solutions are renormalized (under very mild integrability assumptions that are needed in order to make sense of the equations).

Proof of Thm 2.12. From (29) and the integrability properties of the commutator that we have proved, it follows that u_ε is Sobolev in time. We can thus make sense of (26) in the almost everywhere sense. Moreover, we can use the chain rule for Sobolev functions (see, for instance, Section 4.2.2 of [**EG92**]). Thus, we have

$$\partial_t\beta(u_\varepsilon) + b\cdot\nabla\beta(u_\varepsilon) = \beta'(u_\varepsilon)\,(\partial_t u_\varepsilon + b\cdot\nabla u_\varepsilon) = \beta'(u_\varepsilon) r_\varepsilon.$$

We now want to take the limit and recover the distributional formulation. Let us fix $\varphi \in C^\infty_c([0,T)\times\mathbb{R}^d)$, multiply the equation by φ, and integrate. This gives

$$\int_0^T\int_{\mathbb{R}^d} [\partial_t\beta(u_\varepsilon) + b\cdot\nabla\beta(u_\varepsilon)]\,\varphi\, dx\, dt = \int_0^T\int_{\mathbb{R}^d} \beta'(u_\varepsilon)\, r_\varepsilon\,\varphi\, dx\, dt.$$

We now integrate by parts on the left-hand side:

$$-\int_0^T\int_{\mathbb{R}^d} \beta(u_\varepsilon)\,(\partial_t\varphi + b\cdot\nabla\varphi)\; dx\, dt - \int_{\mathbb{R}^d} \beta(u_\varepsilon(0,x))\,\varphi(0,x)\, dx = \int_0^T\int_{\mathbb{R}^d} \beta'(u_\varepsilon)\, r_\varepsilon\,\varphi\, dx\, dt,$$

where we used that b is divergence-free.

Finally, by letting $\varepsilon \to 0$ and using the fact that the commutator r_ε converges strongly to zero in L^1_{loc}, we obtain (up to passing to subsequences) the following equation:

$$-\int_0^T\int_{\mathbb{R}^d} \beta(u)\,(\partial_t\varphi + b\cdot\nabla\varphi)\; dx\, dt - \int_{\mathbb{R}^d} \beta(u(0,x))\,\varphi(0,x)\, dx = 0,$$

which implies that $\beta(u)$ is a distributional solution.

□

3. Connection between PDE and ODE

In this final section, we show how the results on the transport equation are related to the existence and uniqueness of the Regular Lagrangian Flow. The way we present this is based on Ambrosio's breakthrough paper [**Amb04**], where he proved that even BV vector fields have the property that bounded solutions are renormalized . We work with the continuity equation in its general form, i.e., (12).

3.1. Superposition solutions and Superposition Principle. We begin by discussing a fundamental result, the Ambrosio Superposition Principle, which provides a way to describe every solution of the continuity equation, even when a flow may not exist, and under mild assumptions. We will denote Γ_T the space $C([0,T];\mathbb{R}^d)$ of continuous curves in $\mathbb{R}^d$. For every $x \in \mathbb{R}^d$ we consider a probability measure $\eta_x \in \mathcal{P}(\Gamma_T)$ supported on the trajectories $\gamma(t)$ that starts from x at time $t = 0$, i.e $\gamma(0) = x$ and are absolutely continuous solutions of the ODE. On the technical point of view, all families $\{\eta_x\}_{x\in\mathbb{R}^d}$ considered in the following discussions are weakly measurable. That is, for every function $\Phi \in C_b(\Gamma_T)$, the map

$$x \mapsto \langle \eta_x, \Phi\rangle = \int_{\Gamma_T} \Phi(\gamma)\, d\eta_x(\gamma)$$

is measurable.

Definition 2.15 (Superposition solution)**.** Let $\bar\mu \in \mathcal{M}(\mathbb{R}^d)$ be an initial measure, representing the distribution of mass at $t = 0$. The superposition solution induced by the family $\{\eta_x\}_{x\in\mathbb{R}^d}$ is the family of measures $\mu_t^{\eta_x} \in \mathcal{M}(\mathbb{R}^d)$, for $t \in [0,T]$, defined as follows:

$$\langle \mu_t^{\eta_x}, \varphi\rangle = \int_{\mathbb{R}^d} \left(\int_{\Gamma_T} \varphi(\gamma(t))\, d\eta_x(\gamma) \right) d\bar\mu(x) \quad \forall \varphi \in C_c(\mathbb{R}^d). \tag{33}$$

To better understand what is happening, note the following. If the associated vector field has, for each $x \in \mathbb{R}^d$, only one solution to the ODE (5), as discussed, for example, in the previous chapter, then the only possibility for η_x is for it to be $\delta_{X(\cdot,x)}$. In this case, the superposition solution becomes:

$$\langle \mu_t^{\delta_{X(\cdot,x)}}, \varphi\rangle = \int_{\mathbb{R}^d} \left(\int_{\Gamma_T} \varphi(\gamma(t))\, d\delta_{X(\cdot,x)}(\gamma) \right) d\bar\mu(x) = \int_{\mathbb{R}^d} \varphi(X(t,x))\, d\bar\mu(x),$$

which is equation (17).
Superposition solutions can thus be interpreted as the average of the pushforwards over all possible trajectories. We now prove that (33) always gives a solution of the continuity equation (15). Since the measure η_x is concentrated on solutions of the ODE starting from $x \in \mathbb{R}^d$ for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$, using Fubini's theorem, we deduce that

$$\dot\gamma(t) = b(t,\gamma(t)) \quad \text{for} \quad \mathscr{L}^d \otimes \eta_x\text{-a.e. } (x,\gamma) \in \mathbb{R}^d \times \Gamma_T$$

for $\mathscr{L}^1$-a.e. $t \in [0,T]$. Furthermore notice that for any $\varphi \in C_c^\infty(\mathbb{R}^d)$, the map $t \mapsto \langle \mu_t^{\eta_x}, \varphi\rangle$ is Lipschitz. Indeed, we can estimate

$$|\varphi(\gamma(t)) - \varphi(\gamma(s))| \le \|\nabla\varphi\|_\infty \|\gamma(t) - \gamma(s)\| \le \|\nabla\varphi\|_\infty \|b\|_\infty |t-s|,$$

and this implies

$$\begin{aligned} |\langle \mu_t^{\eta_x}, \varphi\rangle - \langle \mu_s^{\eta_x}, \varphi\rangle| &= \left| \int_{\mathbb{R}^d} \int_{\Gamma_T} \left(\varphi(\gamma(t)) - \varphi(\gamma(s))\right) d\eta_x(\gamma)\, d\bar\mu(x) \right| \\ &\le \|\nabla\varphi\|_\infty \|b\|_\infty |t-s| \bar\mu \left(\operatorname{supp}(\varphi) + B_{T\|b\|_\infty}(0) \right). \end{aligned}$$

Hence the distributional derivative of the map $t \mapsto \langle \mu_t^{\eta_x}, \varphi \rangle$ coincides with the pointwise one. We finally compute the pointwise derivative

$$\begin{aligned}\frac{d}{dt}\langle \mu_t^{\eta_x}, \varphi \rangle &= \frac{d}{dt}\int_{\mathbb{R}^d}\left(\int_{\Gamma_T} \varphi(\gamma(t))\, d\eta_x(\gamma)\right) d\bar{\mu}(x) \\ &= \int_{\mathbb{R}^d}\int_{\Gamma_T} \nabla\varphi(\gamma(t)) \cdot b(t,\gamma(t))\, d\eta_x(\gamma)\, dx \\ &= \int_{\mathbb{R}^d} b(t,x) \cdot \nabla\varphi(x)\, d\mu_t(x).\end{aligned}$$

where in the last equality we have used the definition of μ_t. This last equality proves that (33) gives a solution of the (15).

As a corollary of what we have proved, by choosing $\eta_x = \delta_{X(\cdot,x)}$, we obtain the following solution to the continuity equation, when there exists a regular Lagrangian flow.

Proposition 2.16 Let b be a bounded and divergence free vector field, and let $X(t,x)$ be a regular Lagrangian flow associated with it. Then, the Cauchy problem

$$\begin{cases}\partial_t\mu_t + \nabla \cdot (b\mu_t) = 0, \\ \mu_0 = \bar{\mu}\end{cases}$$

has the solution

$$\mu_t = (X_t)_\# \bar{\mu}. \tag{34}$$

Note that it is the generalization of (17) in the low regularity setting.

The so-called Ambrosio Superposition Principle essentially states that, under certain assumptions, the construction in (33) can be reversed—i.e., any solution of the continuity equation can be represented in this form.

Theorem 2.17 (Superposition principle) Let $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded vector field, and let $\mu_t \in \mathcal{M}_+(\mathbb{R}^d)$ be a positive, locally finite, measure-valued solution of the continuity equation. Then μ_t is a superposition solution, i.e., there exists a family $\{\eta_x\}_{x\in\mathbb{R}^d} \subset \mathcal{P}(\Gamma_T)$, with each η_x concentrated on absolutely continuous integral curves of the ODE starting from x, for $\bar{\mu}$-almost every $x \in \mathbb{R}^d$, such that $\mu_t = \mu_t^{\eta_x}$ for every $t \in [0,T]$.

This theorem will be proved in the third subsection.

3.2. Existence and uniqueness of the Regular Lagrangian Flow. We will use the superposition principle to prove the existence of the regular Lagrangian flow under suitable assumptions. Throughout, we assume that the vector field is bounded, divergence-free, and belongs to $L^1_{\text{loc}}(\mathbb{R}, W^{1,p}_{\text{loc}}(\mathbb{R}^d, \mathbb{R}^d))$. Under these assumptions, we have shown that there exists a unique solution, starting from any initial datum $\bar{u} \in L^\infty$, that remains bounded. We will use the notation $\mu_t \in L^\infty([0,T] \times \mathbb{R}^d)$ for measures that are absolutely continuous whose density is essentially uniformly bounded. We will need a preparatory proposition that will be proved in the third subsection.

Proposition 2.18 Let $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded vector field. Assume that the continuity equation (12) has the uniqueness property in $L^\infty([0,T] \times \mathbb{R}^d)$. Consider a family $\{\eta_x\}_{x\in\mathbb{R}^d} \subset P(\Gamma_T)$ such that η_x is concentrated on absolutely continuous integral solutions of the ordinary differential equation starting from x, for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$. Assume that the superposition solution of the continuity equation $\mu_t^{\eta_x}$ induced by this family belongs to $L^\infty([0,T] \times \mathbb{R}^d)$. Then η_x is a Dirac mass for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$.

Assume now by contradiction there exists two different Regular Lagrangian Flows $X(t,x)$, $Y(t,x)$ we will construct a counterexample to proposition 2.18. Consider for a.e $x \in \mathbb{R}^d$ $\eta_x = \frac{1}{2}(\delta_{X(\cdot,x)} + \delta_{Y(\cdot,x)})$. One can see that first, by hypotheses η_x is not a dirac mass. Second, again by hypotheses, since X, Y are regular lagrangian flows, that the measure is concentrated on absolutely continuous integral curves. We finally proves that the induced superposition solution belongs to $L^\infty([0,T]\times\mathbb{R}^d)$. Let $\varphi \in C_c^\infty(\mathbb{R}^d)$,

$$\begin{aligned}\langle \mu_t^{\eta_x}, \varphi\rangle &= \int_{\mathbb{R}^d}\int_{\Gamma_T} \varphi(\gamma(t))\, d\eta_x(\gamma)\, dx \\ &= \int_{\mathbb{R}^d}\int_{\Gamma_T} \varphi(\gamma(t))\, d\left(\tfrac{1}{2}\delta_{X(\cdot,x)} + \tfrac{1}{2}\delta_{Y(\cdot,x)}\right)\, dx \\ &= \int_{\mathbb{R}^d} \tfrac{1}{2}\left(\varphi(X(t,x)) + \varphi(Y(t,x))\right)\, dx \\ &= \tfrac{1}{2}\int_{\mathbb{R}^d} \varphi(y)\, d\left((X(t,\cdot))_{\#}\mathscr{L}^d\right)(y) + \tfrac{1}{2}\int_{\mathbb{R}^d} \varphi(y)\, d\left((Y(t,\cdot))_{\#}\mathscr{L}^d\right)(y).\end{aligned}$$

This means that

$$\mu_t^{\eta_x} = \tfrac{1}{2}\left((X(t,\cdot))_{\#}\mathscr{L}^d + (Y(t,\cdot))_{\#}\mathscr{L}^d\right).$$

That is a bounded solution since the vector field is bounded and this gives a contradiction. We have thus proved uniqueness of the Regular Lagrangian Flow under our assumptions. Finally we prove that one regular Lagrangian flow always exists. Note that in our setting, since the flow is divergence free, there is the non-negative, constant (and bounded) solution $u(t,x) = 1$ or equivalently $\mu_t = \mathscr{L}^d$. By applying Theorem 2.17, we obtain that $\mu_t = \mu_t^{\eta_x}$ for some family $\{\eta_x\}_{x\in\mathbb{R}^d}$, with η_x concentrated on absolutely continuous integral solutions of the ODE starting from x, for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$. But again, by proposition 2.18, η_x is a Dirac mass for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$. We denote $X(\cdot,x)$ for the element in which η_x is concentrated. Clearly the map $X(t,x)$ satisfies the first two properties of Regular Lagrangian Flows, since η_x is supported on integral curves of the vector fields. The last property follows from how η_x was built. In particular for every $\varphi \in C_c^\infty(\mathbb{R}^d)$,

$$\begin{aligned}\langle \mu_t^{\eta_x}, \varphi\rangle &= \int_{\mathbb{R}^d}\int_{\Gamma_T} \varphi(\gamma(t))\, d\eta_x(\gamma)\, dx \\ &= \int_{\mathbb{R}^d} \varphi(X(t,x)dx.\end{aligned}$$

But on the other hand,

$$\langle \mu_t^{\eta_x}, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(x)\, dx,$$

so for every $t \in [0,T]$, we have

$$\int_{\mathbb{R}^d} \varphi(X(t,x))\, dx = \int_{\mathbb{R}^d} \varphi(x)\, dx,$$

which is precisely the last required condition for a Regular Lagrangian Flow.

3.3. Technical proofs. In this subsection, we will prove Theorem 2.17 and Proposition 2.18. We then conclude with a theorem that links uniqueness for the ODE and for the PDE. We begin with the proof of the first result, where the strategy—as is often the case in this thesis—is to regularize the vector field and then pass to the limit to obtain the desired result. For the reader's convenience, we recall the statement of the theorem below. The proof here will substantially follow Theorem 6.22 in [**Cri07**].

Theorem 2.19 Let $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ be a bounded vector field, and let $\mu_t \in \mathcal{M}_+(\mathbb{R}^d)$ be a positive, locally finite, measure-valued solution of the continuity equation. Then μ_t is a superposition solution, i.e., there exists a family $\{\eta_x\}_{x\in\mathbb{R}^d} \subset \mathcal{P}(\Gamma_T)$, with each η_x concentrated on absolutely continuous integral curves of the ODE starting from x, for $\bar{\mu}$-almost every $x \in \mathbb{R}^d$, such that $\mu_t = \mu_t^{\eta_x}$ for every $t \in [0,T]$.

Proof. The proof will consist of several steps. Through all the proof we will need to change variables, so we must first understand how the domain of integration changes. Notice that if $x \in B_R$, and since $\left|\frac{d}{dt}X(t,\cdot)\right| \le \|b\|_\infty$, then $X(t,x) \in B_{R+T\|b\|_\infty}$.

Step 1: Uniform control of the local masses.

The measures μ_t are just locally finite; however, due to the finite speed of propagation of the transport equation (by the assumption that the vector field b is uniformly bounded), it is easy to show the existence of a constant m_R, independent of time, such that

$$\mu_t(B_R(0)) \leq m_R \quad \text{for every } t \in [0,T]. \tag{35}$$

This can be proved by integrating over suitable space-time cones; see, for instance, Lemma 2.12 in [**ABL04**]. The argument presented there for bounded solutions carries over without any essential change to the case of measure-valued solutions.

Step 2: Construction of an adapted convolution kernel.

We want to construct a positive convolution kernel $\rho \in C^k(\mathbb{R}^d)$, with the following properties:

- $\operatorname{spt}\rho = \mathbb{R}^d$,
- $\displaystyle\int_{\mathbb{R}^d} \rho(x)\,dx = 1$,

in such a way that, for some function $\tilde{m}_R$, we have

$$\mu_t^\varepsilon(B_R(0)) \leq \tilde{m}_R \quad \text{for every } t \in [0,T] \text{ and every } \varepsilon \in (0,1), \tag{36}$$

where $\mu_t^\varepsilon := \mu_t * \rho_\varepsilon$, and $\rho_\varepsilon(x) := \varepsilon^{-d}\rho(x/\varepsilon)$. This can be achieved as follows. We define ρ as a weighted sum of standard, non negative, convolution kernels ρ_j supported on $B_j(0)$:

$$\rho(x) = \sum_{j=1}^\infty c_j \rho_j(x)$$

with weight c_j to be chosen. First, in order to have a positive kernel with integral 1, it suffices to require $c_j > 0$ and

$$\sum_{j=1}^\infty c_j = 1.$$

Moreover, in order to have a C^k function, it suffices to require

$$\sum_{j=1}^\infty c_j \|\rho_j\|_{C^k} < \infty$$

which is satisfied for each $k > 0$, provided the sequence goes to zero sufficiently fast. We now focus on (36):

$$\begin{aligned}
\mu_t^\epsilon(B_R(0)) &= \int_{B_R(0)} \mu_t^\epsilon(x)\,dx \\
&= \int_{B_R(0)} \int_{\mathbb{R}^d} \rho^\epsilon(x-y)\,d\mu_t(y)\,dx \\
&= \int_{B_R(0)} \int_{\mathbb{R}^d} \sum_{j=1}^\infty c_j \rho_j^\epsilon(x-y)\,d\mu_t(y)\,dx \\
&= \sum_{j=1}^\infty c_j \int_{B_R(0)} \int_{\mathbb{R}^d} \rho_j^\epsilon(x-y)\,d\mu_t(y)\,dx \\
&= \sum_{j=1}^\infty c_j \int_{\mathbb{R}^d} \int_{B_R(0)} \rho_j^\epsilon(x-y)\,dx\,d\mu_t(y) \\
&= \sum_{j=1}^\infty c_j \int_{B_{R+j}(0)} \int_{B_R(0)} \rho_j^\epsilon(x-y)\,dx\,d\mu_t(y) \\
&\leq \sum_{j=1}^\infty c_j \int_{B_{R+j}(0)} d\mu_t(y) \\
&\leq \sum_{j=1}^\infty c_j m_{R+j}.
\end{aligned}$$

We thus define

$$\tilde{m}_R = \sum_{j=1}^{\infty} c_j m_{R+j} \tag{37}$$

and we want it to be finite. To end this, it suffices to choose c_j such that

$$c_j m_j, c_j m_{j+1}, \dots, c_j m_{2j} \leq \frac{1}{2^j}.$$

Indeed, with this choice, we can estimate

$$\tilde{m}_R = \sum_{j=1}^{\infty} c_j m_{R+j} = \sum_{j=1}^{R-1} c_j m_{R+j} + \sum_{j=R}^{\infty} c_j m_{R+j} \leq \sum_{j=1}^{R-1} c_j m_{R+j} + \sum_{j=R}^{\infty} \frac{1}{2^j} < \infty.$$

Finally, note that if a sequence satisfies (37), then a sequence with lower terms (by comparing term by term) also satisfies it. Therefore, we can construct a sequence that satisfies all the conditions.

Step 3: Smoothing.

We now use the kernel ρ constructed in the previous step, and we consider:

$$\mu_t^\epsilon(x) = (\mu_t * \rho^\epsilon)(x) \quad \text{and} \quad b^\epsilon(t,x) = \frac{(b(t,\cdot)\mu_t) * \rho^\epsilon(x)}{\mu_t^\epsilon(x)}$$

where we use in a crucial way that μ_t and ρ are positive. One can see that b^ϵ is bounded by the bound on b:

$$|b^\epsilon(t,x)| = \frac{|(b(t,\cdot)\mu_t) * \rho^\epsilon(x)|}{\mu_t^\epsilon(x)} \leq \frac{\int_{\mathbb{R}^d} |b(t,y)\rho^\epsilon(x-y)| d\mu_t(y)}{\mu_t^\epsilon(x)} \leq \frac{\|b\|_{L^\infty} \mu_t^\epsilon(x)}{\mu_t^\epsilon(x)} \leq \|b\|_{L^\infty}.$$

Moreover, μ_t^ϵ is a solution of the continuity equation with vector field b^ϵ:

$$\partial_t \mu_t^\epsilon + \operatorname{div}(b^\epsilon \mu_t^\epsilon) = \partial_t \mu_t * \rho^\epsilon + \operatorname{div}(b(t,\cdot)\mu_t) * \rho^\epsilon = 0$$

Since b^ϵ is bounded and smooth in space, there exists a unique flow $X^\epsilon(t,x)$ associated with b^ϵ, which is globally defined for $t \in [0,T]$. Moreover, by the representation formula (17) we have

$$\mu_t^\epsilon = X^\epsilon(t,\cdot)_\# \mu_0^\epsilon.$$

We thus define

$$\eta_x^\epsilon = \delta_{X^\epsilon(\cdot,x)} \in \mathcal{P}(\Gamma_T) \quad \text{and} \quad \eta^\epsilon = \mu_0^\epsilon \otimes \eta_x^\epsilon \in \mathcal{M}_+(\mathbb{R}^d \times \Gamma_T),$$

where the tensor product satisfies

$$\langle \eta^\epsilon, \Phi(x,\gamma) \rangle = \int_{\mathbb{R}^d} \int_{\Gamma_T} \Phi(x,\gamma)\, d\eta_x^\epsilon(\gamma)\, d\mu_0^\epsilon(x)$$

for every function $\Phi \in C_b(\mathbb{R}^d \times \Gamma_T)$ whose support has a compact projection over $\mathbb{R}^d$.

Step 4: Tightness.

In this section we want to find a limit for η^ϵ by using a compactness argument. We consider the functional

$$\Psi : \mathbb{R}^d \times \Gamma_T \to [0, +\infty], \quad \Psi(x,\gamma) = \int_0^T |\dot\gamma(t)|^2\, dt,$$

which takes the value $+\infty$ for all curves $\gamma \in \Gamma_T$ not belonging to $AC^2([0,T];\mathbb{R}^d)$, the space of absolutely continuous maps with square-integrable derivative.

For every $R > 0$, we define the set

$$\Gamma_{T,R} := \{\gamma \in \Gamma_T : \gamma(0) \in B_R(0)\},$$

and the localized functional

$$\Psi_R(x,\gamma) = \begin{cases} \Psi(x,\gamma) & \text{if } (x,\gamma) \in B_R(0) \times \Gamma_{T,R}, \\ +\infty & \text{otherwise.} \end{cases}$$

We will now show that Ψ_R is coercive for every $R > 0$, in particular we will show that every sub-level set is relatively compact. For every finite $\lambda > 0$, let $(x,\gamma) \in \{\Psi_R \leq \lambda\}$. Since λ is finite, we deduce $x \in B_R(0)$ and $\gamma(0) \in B_R(0)$. For every $s, t \in [0, T]$, we can compute the distance between $\gamma(t)$ and $\gamma(s)$ as

$$|\gamma(t) - \gamma(s)| = \left| \int_s^t \gamma'(\tau)\, d\tau \right|.$$

By applying the Cauchy-Schwarz inequality, we get

$$|\gamma(t) - \gamma(s)| \le \left(\int_s^t |\gamma'(\tau)|^2\, d\tau \right)^{1/2} |t-s|^{1/2}$$

Now, since $\Psi_R(x, \gamma) \le \lambda$, we have

$$|\gamma(t) - \gamma(s)| \le \lambda^{1/2} |t-s|^{1/2}.$$

Finally, by applying the Ascoli-Arzelà theorem, since the functions are equibounded and equicontinuous, we deduce that each sub-level set is relatively compact in $\mathbb{R}^d \times \Gamma_T$. We define also the truncated measure $\eta_{\epsilon,R}$ as

$$\eta^{\epsilon,R} = \eta^\epsilon \llcorner (B_R(0) \times \Gamma_{T,R}) \in \mathcal{M}^+(\mathbb{R}^d \times \Gamma_T),$$

which is in fact a finite measure. Now, we want to evaluate the value of the truncated measure $\eta^{\epsilon,R}$ on the localized functional Ψ^R. Using the definition of η_ϵ, the fact that μ_ϵ is a solution of the continuity equation, and the uniform bound on b_ϵ, we can estimate

$$\begin{aligned}
\langle \eta^{\epsilon,R}, \Psi \rangle &= \int_{\mathbb{R}^d \times \Gamma_T} \Psi(x,\gamma)\, d\eta^{\epsilon,R}(x,\gamma) \\
&= \int_{B_R(0)} \int_{\Gamma_{T,R}} \Psi(x,\gamma)\, d\eta_x^\epsilon(\gamma)\, d\mu_0^\epsilon(x) \\
&= \int_{B_R(0)} \int_0^T \left| \frac{\partial X^\epsilon}{\partial t}(t,x) \right|^2 dt\, d\mu_0^\epsilon(x) \\
&= \int_0^T \int_{B_R(0)} |b^\epsilon(t, X^\epsilon(t,x))|^2\, d\mu_0^\epsilon(x)\, dt \\
&= \int_0^T \int_{X^\epsilon(t, B_R(0))} |b^\epsilon(t,x)|^2\, d\mu_t^\epsilon(x)\, dt \\
&\le \int_0^T \int_{B_{R+T\|b\|_\infty}(0)} |b(t,x)|^2\, d\mu_t^\epsilon(x)\, dt \\
&\le \|b\|_{L^\infty}^2 \int_0^T \int_{B_{R+T\|b\|_\infty}(0)} d\mu_t^\epsilon(x)\, dt \\
&\le T\|b\|_{L^\infty}^2 \tilde{m}_{R+T\|b\|_{L^\infty}}.
\end{aligned}$$

Together with the coercivity of Ψ_R, this gives that, for every fixed R, the family

$$\{\eta^{\epsilon,R}\}_\epsilon \subset M^+(\mathbb{R}^d \times \Gamma_T)$$

is tight (see Theorem A.23). Noticing that this family is also equi-bounded by the simple estimate

$$\eta^{\epsilon,R}(\mathbb{R}^d \times \Gamma_T) = \mu_0^\epsilon(B_R(0)) \le \tilde{m}_R,$$

we can apply the Prokhorov theorem (see Theorem A.21) and deduce that it is relatively sequentially narrowly compact. Finally, with a standard diagonal argument, it is possible to construct a sequence $\{\eta^{\epsilon_i}\}_i \subset \mathcal{M}^+(\mathbb{R}^d \times \Gamma_T)$ and a measure $\eta \in M^+(\mathbb{R}^d \times \Gamma_T)$ such that

$$\eta^{\epsilon_i,R} \to \eta^R \text{ narrowly in } \mathcal{M}^+(\mathbb{R}^d \times \Gamma_T) \text{ for every } R.$$

Step 5: Disintegration of the measure.
Up to now, we have constructed a measure $\eta \in \mathcal{M}^+(\mathbb{R}^d \times \Gamma_T)$. It is clear that the marginal of

η on $\mathbb{R}^d$ is $\bar{\mu}$. Indeed, for every $\epsilon \in (0,1)$, we have

$$(\pi_{\mathbb{R}^d})_\# \eta^\epsilon = \mu_0^\epsilon,$$

and the truncated narrow convergence is inherited by the marginal. This, combined with the fact that $\mu_0^\epsilon \rightharpoonup \bar{\mu}$ weakly in $\mathbb{R}^d$, implies that

$$(\pi_{\mathbb{R}^d})_\# \eta = \bar{\mu}.$$

This allows us to apply the disintegration theorem (see Theorem A.24) , obtaining

$$\eta = \bar{\mu} \otimes \eta_x,$$

with $\eta_x \in \mathcal{P}(\Gamma_T)$ for $\bar{\mu}$-a.e. $x \in \mathbb{R}^d$. We now verify that $\mu_t^{\eta_x} = \mu_t$. Indeed, for every $\epsilon \in (0,1)$, applying the definition of η_x^ϵ and using the fact that $\mu_t^\epsilon = X^\epsilon(t,\cdot)_\# \mu_0^\epsilon$, we obtain:

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

for every $\varphi \in C_c(\mathbb{R}^d)$. Moreover we observe that

$$\int_{\mathbb{R}^d} \varphi(x)\, d\mu_t^\epsilon(x) \longrightarrow \int_{\mathbb{R}^d} \varphi(x)\, d\mu_t(x) = \langle \mu_t, \varphi \rangle, \tag{38}$$

by the truncated narrow convergence of $\mu_t^\epsilon \to \mu_t$. On the other hand, we have

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(X^\epsilon(t,x))\, d\mu_0^\epsilon(x) &= \int_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t))\, d\eta^\epsilon(x,\gamma) \\ &\longrightarrow \int_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t))\, d\eta(x,\gamma) \\ &= \int_{\mathbb{R}^d} \int_{\Gamma_T} \varphi(\gamma(t))\, d\eta_x(\gamma)\, d\bar{\mu}(x) = \langle \mu_t^{\eta_x}, \varphi \rangle, \end{aligned} \tag{39}$$

by the truncated narrow convergence $\eta^\epsilon \to \eta$ along the chosen subsequence.

Combining (38) and (39), and recalling from earlier that

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

we finally deduce that:

$$\mu_t^{\eta_x} = \mu_t,$$

as wanted.

Step 6: η is concentrated on solutions of the ODE.

To conclude the proof, we now aim to show that η_x is concentrated on solutions of the ODE for $\bar{\mu}$-almost every $x \in \mathbb{R}^d$. For clarity, we assume that b has essentially bounded support. The argument can be easily adapted to the general case, as all estimates involved are local in nature. It suffices to show that, for every $t \in [0,T]$ and every integer $R > 0$,

$$\int_{B_R(0) \times \Gamma_T} \left| \gamma(t) - x - \int_0^t b(s,\gamma(s))\, ds \right| d\eta(x,\gamma) = 0. \tag{40}$$

Indeed, for every $t \in [0,T]$ and every $R \in \mathbb{Z}_+$, this yields a $\bar{\mu}$-negligible set $N_{t,R} \subset B_R(0)$ such that the identity

$$\gamma(t) = x + \int_0^t b(s,\gamma(s))\, ds \tag{41}$$

holds for every $x \in B_R(0) \setminus N_{t,R}$, for η_x-almost every $\gamma \in \Gamma_T$. Letting

$$N = \bigcup_{t \in [0,T] \cap \mathbb{Q}} \bigcup_{R \in \mathbb{Z}_+} N_{t,R},$$

which is clearly $\mathcal{L}^d$-negligible, and using the continuity of each trajectory γ, we deduce that (41) holds for every $t \in [0,T]$, for every $x \in \mathbb{R}^d \setminus N$, and for η_x-almost every $\gamma \in \Gamma_T$. We end the proof by proving (41) for every $t \in [0,T]$ and every positive integer R. One major issue arises from the fact that the function appearing in (41) is not even continuous, and therefore cannot be used as a test function. Let a be a continuous vector field. We show that

$$\int_{B_R(0)\times\Gamma_T}\left|\gamma(t)-x-\int_0^t a(s,\gamma(s))\,ds\right|\,d\eta(x,\gamma)\le\int_0^T\int_{\mathbb{R}^d}|b(s,x)-a(s,x)|\,d\mu_s(x)\,ds$$

For every $\epsilon\in(0,1)$

$$\begin{aligned}&\int_{B_R(0)\times\Gamma_T}\left|\gamma(t)-x-\int_0^t a(s,\gamma(s))\,ds\right|\,d\eta^\epsilon(x,\gamma)\\&=\int_{B_R(0)}\left|X^\epsilon(t,x)-x-\int_0^t a(s,X^\epsilon(s,x))\,ds\right|\,d\mu_0^\epsilon(x)\\&\le\int_0^t\int_{B_R(0)}|(b^\epsilon-a)(s,X^\epsilon(s,x))|\,d\mu_0^\epsilon(x)\,ds\\&\le\int_0^t\int_{B_R(0)}(|b^\epsilon-a^\epsilon|(s,X^\epsilon(s,x))+|a^\epsilon-a|(s,X^\epsilon(s,x)))\,d\mu_0^\epsilon(x)\,ds\\&\le\int_0^t\int_{B_{R+T\|b\|_\infty}(0)}(|b^\epsilon-a^\epsilon|(s,x)+|a^\epsilon-a|(s,x))\,d\mu_s^\epsilon(x)\,ds,\end{aligned}$$

where we have used the definition of η^ϵ, the identity $\mu_t^\epsilon=X^\epsilon(t,\cdot)_\#\mu_0^\epsilon$, and we have introduced the regularization

$$a^\epsilon(t,x):=\frac{(a(t,\cdot)\mu_t)*\rho^\epsilon(x)}{\mu_t^\epsilon(x)},$$

which corresponds to the standard mollification of the vector field a with respect to the measure μ_t. We estimate the first integral. Setting $d=b-a$, we notice that

$$|b^\epsilon-a^\epsilon|=|d^\epsilon|=\left|\frac{(d(t,\cdot)\mu_t)*\rho^\epsilon}{\mu_t^\epsilon}\right|,$$

hence

$$\begin{aligned}\int_0^t\int_{B_{R+T\|b\|_\infty}(0)}|d^\epsilon(s,x)|\,d\mu_s^\epsilon(x)\,ds&=\int_0^t\int_{B_{R+T\|b\|_\infty}(0)}\left|\frac{(d(s,\cdot)\mu_s)*\rho^\epsilon(x)}{\mu_s^\epsilon(x)}\right|\mu_s^\epsilon(x)\,dx\,ds\\&=\int_0^t\int_{B_{R+T\|b\|_\infty}(0)}|(d(s,\cdot)\mu_s)*\rho^\epsilon(x)|\,dx\,ds\\&\le\int_0^t\int_{\mathbb{R}^d}\int_{B_{R+T\|b\|_\infty}(0)}\rho^\epsilon(x-y)\,dx\,|d(s,y)|\,d\mu_s(y)\,ds\\&\le\int_0^T\int_{\mathbb{R}^d}|d(s,y)|\,d\mu_s(y)\,ds.\end{aligned}$$

Since a is continuous, we have that $a^\epsilon\to a$ locally uniformly, hence the second integral vanishes as $\epsilon\to 0$. Passing to the limit in the inequalities above (along the subsequence for which we have truncated narrow convergence of η^ϵ to η), we obtain the desired equality for a, i.e.,

$$\int_{B_R(0)\times\Gamma_T}\left|\gamma(t)-x-\int_0^t a(s,\gamma(s))\,ds\right|\,d\eta(x,\gamma)\le\int_0^T\int_{\mathbb{R}^d}|b(s,x)-a(s,x)|\,d\mu_s(x)\,ds.$$

We now see how to derive (40) from this last equation. By approximation, we chose a sequence $\{a_k\}$ of continuous vector fields such that

$$\int_0^T\int_{\mathbb{R}^d}|b(s,x)-a_k(s,x)|\,d\mu_s(x)\,ds\to 0\quad\text{as}\quad k\to\infty.$$

We can then finally estimate by using the previous two equations and the fact that $\mu_t = \mu_t^{\eta_x}$ to get that

$$\begin{aligned}\int_{B_R(0)\times\Gamma_T} &\left|\gamma(t) - x - \int_0^t b(s,\gamma(s))\,ds\right| d\eta(x,\gamma)\\ &\le \int_{B_R(0)\times\Gamma_T}\left|\gamma(t) - x - \int_0^t a_k(s,\gamma(s))\,ds\right| d\eta(x,\gamma)\\ &\quad + \int_{B_R(0)\times\Gamma_T}\left|\int_0^t b(s,\gamma(s)) - a_k(s,\gamma(s))\,ds\right| d\eta(x,\gamma)\\ &\le 2\int_0^T\int_{\mathbb{R}^d}|b(s,x) - a_k(s,x)|\,d\mu_s(x)\,ds,\end{aligned}$$

and this last term as $k \to \infty$ tends to zero as wanted. Hence, we have proved (40) and we have concluded the proof of the whole theorem. □

Before proving Proposition 2.18 we need the following lemma.

Lemma 2.20 Let $\{\eta_x\}_{x\in\mathbb{R}^d} \subset \mathcal{P}(\Gamma_T)$ be a family of probability measures such that for every $t \in [0,T]$ and every pair of disjoint Borel sets $E_1, E_2 \subset \mathbb{R}^d$, one has

$$\eta_x(\{\gamma : \gamma(t) \in E_1\}) \cdot \eta_x(\{\gamma : \gamma(t) \in E_2\}) = 0 \quad \text{for } \mathcal{L}^d\text{-a.e. } x \in \mathbb{R}^d.$$

Then η_x is a Dirac mass for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$.

Proof. For each $t \in [0,T]$, and for each $x \in \mathbb{R}^d$, we define the pushforward measure $\nu_x^t := (e_t)_\#\eta_x \in \mathcal{P}(\mathbb{R}^d)$, where $e_t : \Gamma_T \to \mathbb{R}^d$ is the evaluation map at time t, given by $e_t(\gamma) := \gamma(t)$.

By assumption, for every pair of disjoint Borel sets $E_1, E_2 \subset \mathbb{R}^d$, we have

$$\nu_x^t(E_1)\cdot\nu_x^t(E_2) = \eta_x(\{\gamma : \gamma(t) \in E_1\}) \cdot \eta_x(\{\gamma : \gamma(t) \in E_2\}) = 0 \quad \text{for } \mathscr{L}^d\text{-a.e. } x.$$

This implies that for $\mathscr{L}^d$-almost every x, the measure ν_x^t is a Dirac mass: indeed, if ν_x^t were not a Dirac mass, we could find two disjoint Borel sets E_1, E_2 with $\nu_x^t(E_1), \nu_x^t(E_2) > 0$, which would contradict the assumption.

Since for almost every $x \in \mathbb{R}^d$, the measure ν_x^t is a Dirac mass, there exists a measurable function $y_t : \mathbb{R}^d \to \mathbb{R}^d$ such that $\nu_x^t = \delta_{y_t(x)}$.

Thus, for $\mathscr{L}^d$-almost every x, the measure η_x is supported on the set

$$\{\gamma \in \Gamma_T : \gamma(t) = y_t(x)\}.$$

This property holds for all $t \in [0,T]$, and in particular for every rational time $t \in [0,T] \cap \mathbb{Q}$. We now define the set

$$\Gamma_x := \bigcap_{t\in[0,T]\cap\mathbb{Q}} \{\gamma \in \Gamma_T : \gamma(t) = y_t(x)\}.$$

Since $\eta_x(\Gamma_x) = 1$, and given that the trajectories $\gamma \in \Gamma_T$ are continuous, the values of γ at rational times uniquely determine the entire curve. Thus, Γ_x consists of a single curve, denoted γ_x, and we conclude that $\eta_x = \delta_{\gamma_x}$.

This shows that for $\mathscr{L}^d$-almost every $x \in \mathbb{R}^d$, η_x is a Dirac mass. □

From the previous lemma, Proposition 2.18 follows naturally: we will construct two bounded solutions of the continuity equation with the same initial datum but different evolutions at a later time, which contradicts uniqueness. For the reader's convenience, we recall the statement of the proposition below.

Proposition 2.21 Let $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded vector field. Assume that the continuity equation (12) has the uniqueness property in $L^\infty([0,T] \times \mathbb{R}^d)$. Consider a family $\{\eta_x\}_{x\in\mathbb{R}^d} \subset P(\Gamma_T)$ such that η_x is concentrated on absolutely continuous integral solutions of the ordinary differential equation starting from x, for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$. Assume that the superposition solution of the continuity equation $\mu_t^{\eta_x}$ induced by this family belongs to $L^\infty([0,T] \times \mathbb{R}^d)$. Then η_x is a Dirac mass for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$.

PROOF. We argue by contradiction, using Lemma 2.20. In particular there exist a time $t_1 \in (0,T]$, a Borel set $C \subset \mathbb{R}^d$ of positive measure, and two disjoint Borel sets $E_1, E_2 \subset \mathbb{R}^d$ such that

$$\eta_x \{\gamma : \gamma(t_1) \in E_1\} \cdot \eta_x \{\gamma : \gamma(t_1) \in E_2\} \neq 0 \quad \text{for every } x \in C.$$

Moreover,by passing to a smaller set C still having strictly positive Lebesgue measure, we can assume that

$$0 < \eta_x \{\gamma : \gamma(t_1) \in E_1\} \leq M\eta_x \{\gamma : \gamma(t_1) \in E_2\} \quad \text{for every } x \in C, \tag{42}$$

for some constant $M > 0$. We now focus on the integral curves that start in C and arrive in E_1 and E_2 at time t_1. In particular we define Define the measures

$$\eta_x^1 := \chi_C(x) \cdot \eta_x \llcorner \{\gamma : \gamma(t_1) \in E_1\} \quad \text{and} \quad \eta_x^2 := M \cdot \chi_C(x) \cdot \eta_x \llcorner \{\gamma : \gamma(t_1) \in E_2\},$$

together with μ_1, μ_2 superposition solutions of the continuity equation induced by η_x^1 and η_x^2 respectively. Notice moreover that the superposition solutions are well-defined even if η_x^1 and η_x^2 are not probability measures, but merely positive measures with finite mass for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$. Fix now $\varphi \in C_c^\infty(\mathbb{R}^d)$, we compute

$$\begin{aligned}
\langle \mu_0^1, \varphi \rangle &= \int_{\mathbb{R}^d} \int_{\Gamma_T} \varphi(\gamma(0))\, d\left(\mathbf{1}_C(x) \cdot \eta_x \llcorner \{\gamma : \gamma(t_1) \in E_1\}\right)(\gamma)\, dx \\
&= \int_C \int_{\{\gamma : \gamma(t_1) \in E_1\}} \varphi(\gamma(0))\, d\eta_x(\gamma)\, dx \\
&= \int_C \varphi(x) \cdot \eta_x \{\gamma : \gamma(t_1) \in E_1\}\, dx
\end{aligned}$$

from which we deduce

$$\mu_0^1 = \eta_x \{\gamma : \gamma(t_1) \in E_1\}\, \mathscr{L}^d \llcorner C.$$

An analogous computation gives

$$\mu_0^2 = M \cdot \eta_x \{\gamma : \gamma(t_1) \in E_2\}\, \mathscr{L}^d \llcorner C.$$

By (42) we have that

$$\mu_0^1 \leq \mu_0^2 \quad \text{as measures on } \mathbb{R}^d.$$

Thus μ_0^1 is absolutely continuous with respect to μ_0^2, hence there exists f such that $\mu_0^1 = f\mu_0^2$. Set

$$\tilde{\eta}_x^2 = f(x)M \cdot \mathbf{1}_C(x) \cdot \eta_x \llcorner \{\gamma : \gamma(t_1) \in E_2\}$$

and let $\tilde{\mu}_t^2$ the superposition solution (until time t_1) induced by $\tilde{\eta}$. By construction $\mu_0^1 = \tilde{\mu}_0^2$, moreover one can see

$$\langle \mu_{t_1}^1, \varphi \rangle = \int_C \int_{\{\gamma : \gamma(t_1) \in E_1\}} \varphi(\gamma(t_0))\, d\eta_x(\gamma)\, dx.$$

and

$$\langle \tilde{\mu}_{t_1}^2, \varphi \rangle = \int_C f(x)M \int_{\{\gamma : \gamma(t_1) \in E_2\}} \varphi(\gamma(t_1))\, d\eta_x(\gamma)\, dx.$$

From the last two equations, we can see that $\mu_{t_1}^1$ is supported on E_1, while $\tilde{\mu}_{t_1}^2$ is supported on E_2. In particular, μ_t^1 and $\tilde{\mu}_t^2$ are two solutions in $L^\infty([0,t_1] \times \mathbb{R}^d)$ of the continuity equation with the same initial datum, but they are different at time $t = t_1$, violating the uniqueness assumption.

□

We conclude this chapter with the following theorem, which highlights a criterion for the uniqueness of solutions to the ODE.

Theorem 2.22 Let $A \subset \mathbb{R}^d$ be a Borel set. The following two properties are equivalent for a bounded vector field b:

(a) Solutions of the ODE are unique for any $x \in A$.
(b) Positive measure-valued solutions of the PDE are unique for any $\bar{\mu}$ concentrated in A, i.e., such that $\bar{\mu}(\mathbb{R}^d \setminus A) = 0$.

Proof. **(a) $\implies$ (b)** Let μ_t be a positive measure-valued solution of the continuity equation with initial data $\bar{\mu}$. Applying Theorem 2.17, we deduce that $\mu_t = \mu_t^{\eta_x}$, with $\eta_x \in P(\Gamma_T)$ concentrated on the absolutely continuous integral solutions of the ODE starting from x, for every point $x \in A$. But assumption (a) precisely means that, for every $x \in A$, the solution is unique. Hence, for every $x \in A$, the measure η_x is a Dirac mass supported on the unique trajectory starting from x, and eventually this gives an explicit formula for the solution μ_t, which is therefore unique.
(b) $\implies$ (a) Assume by contradiction that for some $x \in A$ there exist two different solutions $\gamma(t)$ and $\tilde{\gamma}(t)$ of the ODE starting from x. Then consider $\mu_t = \delta_{\gamma(t)}$ and $\tilde{\mu}_t = \delta_{\tilde{\gamma}(t)}$. We clearly have $\mu_0 = \tilde{\mu}_0 = \delta_x$. It is to verify that μ_t and $\tilde{\mu}_t$ are solutions of the continuity equation, but since they are different, we are violating assumption (b). □

The applicability of this theorem is limited by the fact that, on the one hand, pointwise uniqueness properties for the ODE are known only in special cases, such as when there is a Lipschitz or Osgood condition on b. On the other hand, uniqueness for general measure-valued solutions is also known only in special situations.

CHAPTER 3

Quantitative estimates for the Regular Lagrangian Flow

1. Introduction

In this chapter, we illustrate several ideas related to ODEs in the Sobolev setting. Rather than relying on the transport equation, this approach focuses on establishing *a priori* estimates for the Regular Lagrangian Flow itself, which are then used to prove existence, uniqueness, and to investigate its regularity. These ideas originated in the 2008 paper by Crippa and De Lellis [**CL08**], and have since been further developed in several works, including [**BN20, BN21**], where the authors establish sharp regularity results for the Regular Lagrangian Flow.

Another natural problem is the following: if a sequence of vector fields converges to some limit (in a suitable sense), what can be said about the convergence of their associated Regular Lagrangian Flows?

We conclude this section by recalling a key inequality: if a function $f \in W^{1,p}(\mathbb{R}^d)$, then for every $x, y \in \mathbb{R}^d$, the following estimate holds:

$$|f(x) - f(y)| \leq C_d |x-y| \left(M|\nabla f|(x) + M|\nabla f|(y) \right), \tag{43}$$

where $M|\nabla f|$ denotes the Hardy–Littlewood maximal function of the gradient of f, and C_d is a constant depending only on the dimension. This inequality will play a crucial role in all our estimates. For a proof see proposition A.29.

Moreover, we recall a basic fact from harmonic analysis: if $g \in L^p$ for some $p > 1$, then $Mg \in L^p$ as well. This property fails when $p = 1$, in which case the maximal function belongs to a weaker space. For further details, see the appendix on harmonic analysis tools. From now on, we adopt the notation $A \lesssim B$ to indicate the existence of a constant $C > 0$ such that

$$A \leq CB,$$

where C may depend on fixed parameters, such as the dimension of the space. Whenever this dependence is not clear or important, it will be explicitly indicated.

2. Fundamental *a priori* estimate for $p > 1$

2.1. Setup. In this section, we derive a fundamental *a priori* estimate for the Regular Lagrangian Flow. The setting is the following: b_1 and b_2 are two bounded, divergence-free vector fields in $L^1_{\mathrm{loc}}(\mathbb{R}, W^{1,p}(\mathbb{R}^d, \mathbb{R}^d))$ for some $p > 1$, and X_1 and X_2 are the corresponding Regular Lagrangian Flows.

Fix $R > 0$ and $\delta > 0$. We denote by B_r the ball centered at 0 with radius r. Our goal is to give an upper bound for the quantity

$$\Phi_\delta(t) := \int_{B_R} \log\left(1 + \frac{|X_1(t,x) - X_2(t,x)|}{\delta}\right) dx,$$

where t ranges over a fixed interval $[0, T]$.

The motivation for considering this quantity comes from the formal computation

$$\frac{d}{dt} \log(|\nabla X|) \leq \frac{\left|\frac{d}{dt}\nabla X\right|}{|\nabla X|} \leq |\nabla b|(X),$$

which suggests that a bound on the vector field yields some control over the flow. For instance, if the gradient is bounded (as in the Lipschitz case), then the flow is also Lipschitz.

Moreover, any upper bound on Φ_δ provides information about $|X_1 - X_2|$. Indeed, fixing $\gamma > 0$, we have

$$\Phi_\delta(t) \geq \int_{B_R \cap \{|X_1 - X_2| \geq \gamma\}} \log\left(1 + \frac{\gamma}{\delta}\right) = \log\left(1 + \frac{\gamma}{\delta}\right) \mathscr{L}^d \left(B_R \cap \{|X_1 - X_2| \geq \gamma\}\right). \tag{44}$$

We use the notation $b^t(x) := b(t,x)$, and for simplicity, write X_1 for $X_1(t,x)$. To proceed, we compute the derivative of Φ_δ.

$$\begin{aligned}
\Phi'_\delta(t) &\leq \int_{B_R} \frac{\left|\frac{d}{dt}X_1 - \frac{d}{dt}X_2\right|}{\delta + |X_1 - X_2|}\,dx \\
&= \int_{B_R} \frac{|b_1^t(X_1) - b_2^t(X_2)|}{\delta + |X_1 - X_2|}\,dx \\
&\leq \int_{B_R} \frac{|b_1^t(X_1) - b_1^t(X_2) + b_1^t(X_2) - b_2^t(X_2)|}{\delta + |X_1 - X_2|}\,dx \\
&\leq \int_{B_R} \frac{|b_1^t(X_1) - b_1^t(X_2)|}{\delta + |X_1 - X_2|}\,dx + \int_{B_R} \frac{|b_1^t(X_2) - b_2^t(X_2)|}{\delta + |X_1 - X_2|}\,dx \\
&=: (I) + (II).
\end{aligned}$$

We estimate the two terms differently. We will need to change variables, so we must first understand how the domain of integration changes. Notice that if $x \in B_R$, and since $\left|\frac{d}{dt}X(t,\cdot)\right| \leq \|b\|_\infty$, then $X(t,x) \in B_{R+T\|b\|_\infty}$. For the first term we use (43):

$$\begin{aligned}
(I) &\leq \int_{B_R} \frac{|X_1 - X_2|\left(M|\nabla b_1^t|(X_1) + M|\nabla b_1^t|(X_2)\right)}{\delta + |X_1 - X_2|}\,dx \\
&\leq \int_{B_R} \left(M|\nabla b_1^t|(X_1) + M|\nabla b_1^t|(X_2)\right)\,dx \\
&\leq \int_{B_{R+T\|b_1\|_\infty}} M|\nabla b_1^t|(y)\,dy + \int_{B_{R+T\|b_2\|_\infty}} M|\nabla b_1^t|(y)\,dy \\
&\leq C_1 ||\nabla b_1^t||_{L^p}
\end{aligned}$$

where the constant C_1 depends on $R, \|b_1\|_\infty, \|b_2\|_\infty, d, T$.

We have used the fact that the flow preserves the Lebesgue measure, Hölder's inequality, and the boundedness of the maximal function from L^p to itself (for $p > 1$).

Now we estimate the other term:

$$(II) \leq \int_{B_R} \frac{|b_1^t(X_2) - b_2^t(X_2)|}{\delta}\,dx \leq \frac{1}{\delta}\|b_1^t - b_2^t\|_{L^1(B_{R+T\|b_2\|_\infty})}.$$

Putting everything together and integrating in time (and noticing that $\Phi_\delta(0) = 0$), we obtain:

$$\Phi_\delta(t) \leq \frac{1}{\delta}\|b_1 - b_2\|_{L^1_t(L^1(B_{R+T\|b_2\|_\infty}))} + C_1\|\nabla b_1\|_{L^1_t(L^p_x)}. \tag{45}$$

2.2. Uniqueness. Combining (44) and (45) in the case $b = b_1 = b_2$, we obtain

$$\log\left(1 + \frac{\gamma}{\delta}\right) \mathscr{L}^d \left(B_R \cap \{|X_1 - X_2| \geq \gamma\}\right) \leq C_1 \|\nabla b\|_{L^1_t(L^p_x)}.$$

Therefore,

$$\mathscr{L}^d \left(B_R \cap \{|X_1 - X_2| \geq \gamma\}\right) \leq \frac{C_1 \|\nabla b\|_{L^1_t(L^p_x)}}{\log\left(1 + \frac{\gamma}{\delta}\right)}.$$

Letting $\delta \to 0$, we conclude uniqueness of the flow, due to the arbitrariness of R and γ.

2.3. Stability and compactness. It is natural to ask whether the definition of Regular Lagrangian Flow is stable under convergence, and whether a suitable compactness property holds for the flows. In this subsection, we prove both of these results.

Let $b \in L^1_{\text{loc}}(\mathbb{R}, W^{1,p}(\mathbb{R}^d, \mathbb{R}^d))$ be a bounded, divergence-free vector field. Suppose there exists a sequence of divergence-free vector fields (b_n), with associated Regular Lagrangian Flows (X_n), such that

$$b_n \to b \quad \text{in } L^1_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d),$$

and the sequence (b_n) is uniformly bounded in $L^\infty(\mathbb{R} \times \mathbb{R}^d)$. We apply estimates (44) and (45) with $b_1 = b$, $b_2 = b_n$, and obtain:

$$\begin{aligned}\mathscr{L}^d\left(B_R \cap \{|X_n - X| \geq \gamma\}\right) \leq &\frac{1}{\delta \log\left(1 + \frac{\gamma}{\delta}\right)} \|b_n - b\|_{L^1_t(L^1(B_{R+T\|b_n\|_\infty}))} \\ &+ \frac{C_1}{\log\left(1 + \frac{\gamma}{\delta}\right)} \|\nabla b\|_{L^1_t(L^p_x)} =: \mathrm{I} + \mathrm{II}.\end{aligned}$$

Given $\gamma, \eta > 0$, we choose $\delta > 0$ sufficiently small so that $\mathrm{II} \leq \eta/2$. Since $b_n \to b$ in L^1_{loc}, we can then find $\bar{n} \in \mathbb{N}$ such that $\mathrm{I} \leq \eta/2$ for all $n \geq \bar{n}$.

These choices imply that for any $\gamma > 0$ and $R > 0$, and for every $\eta > 0$, there exists $\bar{n} \in \mathbb{N}$ such that

$$\mathscr{L}^d\left(B_R \cap \{|X_n(t,\cdot) - X(t,\cdot)| > \gamma\}\right) \leq \eta \quad \text{for all } n \geq \bar{n}.$$

In particular, this condition implies that X_n converges locally in measure to X. Moreover, since $\left|\frac{d}{dt} X(t,\cdot)\right| \leq \|b\|_\infty$, we have that if $x \in B_R$, then

$$X(t,x) \in B_{R+T\|b\|_\infty}.$$

Therefore, the flows X_n and X are locally uniformly bounded, since the vector fields are uniformly bounded. Thanks to this local uniform boundedness, convergence in measure on bounded sets implies convergence in L^1_{loc}. That is, given $R > 0$ and $\eta > 0$, there exists $\bar{n} \in \mathbb{N}$ such that

$$\int_{B_R} |X_n(t,x) - X(t,x)|\, dx \leq \eta \quad \text{for all } n \geq \bar{n}.$$

In particular this proves that the flows are stable under L^1_{loc} convergence.
Proving compactness follows a similar strategy. Let (b_n) be a sequence of divergence-free vector fields, equibounded in L^∞ and in $L^1_t(W^{1,p}_x)$, and assume $b_n \to b$ in L^1_{loc}. Applying (44) and (45) with $b_1 = b_n$, $b_2 = b_m$, we obtain

$$\begin{aligned}\mathscr{L}^d\left(B_R \cap \{|X_n - X_m| \geq \gamma\}\right) \leq &\frac{1}{\delta \log\left(1 + \frac{\gamma}{\delta}\right)} \|b_n - b_m\|_{L^1_t(L^1(B_{R+T\|b_n\|_\infty}))} \\ &+ \frac{C_1}{\log\left(1 + \frac{\gamma}{\delta}\right)} \|\nabla b_n\|_{L^1_t(L^p_x)} =: \mathrm{I} + \mathrm{II}.\end{aligned}$$

Given $\gamma, \eta > 0$, we first choose $\delta > 0$ small enough so that $\mathrm{II} \leq \eta/2$, using the uniform bound on $\|\nabla b_n\|_{L^1_t(L^p_x)}$. Then, by the strong convergence $b_n \to b$ in L^1_{loc}, we find $\bar{n} \in \mathbb{N}$ such that for all $n, m \geq \bar{n}$, we have $\mathrm{I} \leq \eta/2$. This implies

$$\mathscr{L}^d\left(B_R \cap \{|X_n - X_m| \geq \gamma\}\right) \leq \eta.$$

As before, since the flows X_n and X are locally uniformly bounded, being a Cauchy sequence with respect to (local) convergence in measure implies that they also form a Cauchy sequence in L^1_{loc}. In particular, given $R > 0$ and $\eta > 0$, there exists $\bar{n} \in \mathbb{N}$ such that

$$\int_{B_R} |X_n(t,x) - X_m(t,x)|\, dx \leq \eta \quad \text{for all } n, m \geq \bar{n}. \tag{46}$$

2.4. Existence. We now show another proof of the existence of the flow via a smoothing argument using compactness. Fix a spatial mollifier ρ_ε and define the regularized vector field $b_\varepsilon := b * \rho_\varepsilon$. It is easy to check that

$$\|b_\varepsilon\|_{L^\infty} \leq \|b\|_{L^\infty},$$

and that $b_\varepsilon \to b$ in $L^1_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d)$ and it is equibounded in

$$L^1_{\text{loc}}(\mathbb{R}, W^{1,p}_{\text{loc}}(\mathbb{R}^d, \mathbb{R}^d)).$$

Moreover, it is easy to see that

$$\operatorname{div}(b_\varepsilon) = (\operatorname{div} b) * \rho_\varepsilon,$$

so the vector fields b_ε are divergence-free, since $\operatorname{div} b = 0$ and convolution preserves this property. We are now in a position to apply the compactness result established in the previous subsection to the flows associated to b_ε. In particular, from (46), we have that the sequence of flows converges, up to a subsequence, to a flow X in L^1_{loc}. Finally, it is easy to see that this limit X is indeed a Regular Lagrangian flow for b.

3. Theory for $p = 1$

In this section, we modify the proof from the previous section to handle the case where $b \in W^{1,1}_{\text{loc}}(\mathbb{R}^d, \mathbb{R}^d)$. The proof we present is a slight improvement on the one in [**BC13**], where the authors address this case, as well as other cases not included in the $p > 1$ theory. The key idea is that some of the estimates in the fundamental estimate can be improved via interpolation. Specifically, at certain points, we bound a denominator consisting of the sum of two terms by just one of the terms (for example, $\frac{1}{\delta+|X_1-X_2|} \leq \frac{1}{|X_1-X_2|}$).

We now recall the definition of the weak Lebesgue space:

$$\|f\|_{L^{1,\infty}} := \sup_{\lambda>0} \lambda \mathscr{L}^d \left(\{x \in \mathbb{R}^d : |f(x)| > \lambda\} \right).$$

It is known that the maximal function is not an operator from L^1 to L^1, but rather from L^1 to $L^{1,\infty}$. This means that

$$\|Mf\|_{L^{1,\infty}} \lesssim \|f\|_{L^1}$$

that can be rewritten as

$$\mathscr{L}^d \left(\{x \in \mathbb{R}^d : |f(x)| > \lambda\} \right) \lesssim \frac{\|f\|_{L^1}}{\lambda}.$$

We will need the following interpolation inequality, which shows that if f is bounded and belongs to the weak space $L^{1,\infty}$, then f also belongs to L^1, provided it has compact support:

$$\|f\|_{L^1} \leq \|f\|_{L^{1,\infty}} \left(1 + \log \left(\frac{C\|f\|_{L^\infty}}{\|f\|_{L^{1,\infty}}} \right) \right), \tag{47}$$

where $C = \mathscr{L}^d(\operatorname{supp} f)$ denotes the Lebesgue measure of the support of f. For a proof of this inequality see proposition A.28.

We start as in the fundamental estimate, using the same setup and the same decomposition into (I) and (II), namely:

$$\begin{aligned}
\Phi'_\delta(t) &\leq \int_{B_R} \frac{\left|\frac{d}{dt}X_1 - \frac{d}{dt}X_2\right|}{\delta + |X_1 - X_2|} \, dx \\
&= \int_{B_R} \frac{|b_1^t(X_1) - b_2^t(X_2)|}{\delta + |X_1 - X_2|} \, dx \\
&\leq \int_{B_R} \frac{|b_1^t(X_1) - b_1^t(X_2) + b_1^t(X_2) - b_2^t(X_2)|}{\delta + |X_1 - X_2|} \, dx \\
&\leq \int_{B_R} \frac{|b_1^t(X_1) - b_1^t(X_2)|}{\delta + |X_1 - X_2|} \, dx + \int_{B_R} \frac{|b_1^t(X_2) - b_2^t(X_2)|}{\delta + |X_1 - X_2|} \, dx \\
&=: (I) + (II).
\end{aligned}$$

We estimate (II) as before:

$$(II) \le \int_{B_R} \frac{|b_1^t(X_2) - b_2^t(X_2)|}{\delta}\, dx \le \frac{1}{\delta}\|b_1^t - b_2^t\|_{L^1(B_{R+T\|b_2\|_\infty})}.$$

To estimate (I), we proceed similarly to the previous proof, but we now use the factor δ in the denominator to compensate for the fact that the maximal function belongs only to a weak space. In a similar way to the proof of the $p>1$ case, we get that

$$\begin{aligned}\Phi_\delta'(t) &\lesssim \int_{B_R} \min\left(\frac{2\|b_1\|_\infty}{\delta},\, \left(M|\nabla b_1^t|(X_1) + M|\nabla b_1^t|(X_2)\right)\right) dx\\ &\lesssim \int_{B_{R+T\|b_1\|_\infty}} \min\left(\frac{\|b_1\|_\infty}{\delta},\, M|\nabla b_1^t|\right) dx.\end{aligned}$$

By integrating in time we get that

$$\Phi_\delta(t) \lesssim \frac{1}{\delta}\|b_1 - b_2\|_{L^1_t(L^1(B_{R+T\|b_2\|_\infty}))} + \int_0^T \int_{B_{R+T\|b_1\|_\infty}} \min\left(\frac{\|b_1\|_\infty}{\delta},\, M|\nabla b_1^t|\right) dx\, dt. \tag{48}$$

We focus on the second term. In particular, we define

$$\phi(s,x) := \min\left(\frac{\|b_1\|_\infty}{\delta},\, M|\nabla b_1^s|(x)\right)\chi_{B_{R+\|b_1\|_\infty T}}(x), \tag{49}$$

where $\chi_{B_{R+\|b_1\|_\infty T}}$ denotes the characteristic function of the ball $B_{R+\|b_1\|_\infty T}$.

At this point, it is important to notice that we are not using the distinction between $W^{1,1}$ and BV functions. Fix $\varepsilon > 0$, we will decompose ∇b_1 in the following way:

$$\nabla b_1 = g_\varepsilon + h_\varepsilon$$

With

$$\|g_\varepsilon\|_{L^1_t(L^1_x)} \le \varepsilon,$$

and

$$\|h_\varepsilon\|_{L^1_t(L^2_x)} \le C_\varepsilon$$

where C_ε blows up as $\varepsilon \to 0$ and depends on the integrability of ∇b_1. One can see that by choosing λ in an appropriate way, $g_\varepsilon = \nabla b_1 \cdot \chi_{\{\nabla b > \lambda\}}$ and $h_\varepsilon = \nabla b_1 \cdot \chi_{\{\nabla b \le \lambda\}}$ satisfy the hypotheses.

With this decomposition we have

$$\begin{aligned}\phi(s,x) &= \min\left(\frac{\|b_1\|_\infty}{\delta},\, M|\nabla b_1^s|(x)\right)\chi_{B_{R+\|b_1\|_\infty T}}(x)\\ &= \min\left(\frac{\|b_1\|_\infty}{\delta},\, M\left(|g_\varepsilon^s| + |h_\varepsilon^s|\right)(x)\right)\chi_{B_{R+\|b_1\|_\infty T}}(x)\\ &\le \min\left(\frac{\|b_1\|_\infty}{\delta},\, M|g_\varepsilon^s|(x)\right)\chi_{B_{R+\|b_1\|_\infty T}}(x) + \min\left(\frac{\|b_1\|_\infty}{\delta},\, M|h_\varepsilon^s|(x)\right)\chi_{B_{R+\|b_1\|_\infty T}}(x)\\ &=: \phi_1(s,x) + \phi_2(s,x).\end{aligned}$$

We now want to deal with ϕ_1 using interpolation and with ϕ_2 as in the $p>1$ case.

Note that $\phi_1(s,x)$ have compact support contained in $[0,T]\times B_{R+\|b_1\|_\infty T}$, therefore we want to apply (47). Note that

$$\|\phi_1\|_{L^\infty_{t,x}} \le \frac{\|b_1\|_\infty}{\delta},$$

and

$$\|\phi_1\|_{L^{1,\infty}_{t,x}} \le \|Mg_\varepsilon\|_{L^{1,\infty}_{t,x}} \lesssim \int_0^T \|Mg_\varepsilon^t\|_{L^{1,\infty}_x} dt \lesssim \int_0^T \|g_\varepsilon^t\|_{L^1_x} dt = \|g_\varepsilon\|_{L^1_t(L^1_x)} \lesssim \varepsilon.$$

On the other hand, for ϕ_2, we have, by using the Hardy-Littlewood Maximal inequality for $p=2$,

$$\|\phi_2\|_{L^1_t(L^2_x)} \le \|Mh_\varepsilon\|_{L^1_t(L^2_x)} \lesssim \|h_\varepsilon\|_{L^1_t(L^2_x)} \lesssim C_\varepsilon. \tag{50}$$

We now estimate the integral involving ϕ. We will use (47) for ϕ_1 and (50) for ϕ_2. The second term is easy; as in the proof for $p > 1$, we use Hölder's inequality and obtain

$$\int_0^T \int_{B_{R+T\|b_1\|_\infty}} \phi_2(t,x)\,dx\,dt \lesssim C_\varepsilon.$$

On the other term, we apply (47) to ϕ_1. Notice a couple of things: the measure of the support of ϕ_1 is controlled by a quantity that depends on $T, R, \|b_1\|_{L^\infty}$, which we will call C. Moreover, notice that $z \mapsto \log z$ is increasing, and that the function

$$z \mapsto z\left(1+\log\left(\frac{c}{\delta z}\right)\right)$$

is also increasing for $z \le \frac{c}{\delta}$. Putting all this information together, we get that

$$\int_0^T \int_{B_{R+T\|b_1\|_\infty}} \phi_1(t,x)\,dx\,dt \lesssim \varepsilon\left(1+\log\left(\frac{C\|b_1\|_{L^\infty}}{\delta\varepsilon}\right)\right).$$

Finally, we see that all these inequalities give

$$\Phi_\delta(t) \lesssim \frac{1}{\delta}\|b_1-b_2\|_{L^1_t(L^1(B_{R+T\|b_2\|_\infty}))} + C_\varepsilon + \varepsilon\left(1+\log\left(\frac{C\|b_1\|_{L^\infty}}{\delta\varepsilon}\right)\right), \tag{51}$$

where $\lesssim$ and C depend on $R, \|b_1\|_{L^\infty}, T$.

As in the case $p > 1$, we derive from (51) the analogous version of (44), i.e.

$$\begin{aligned}\mathscr{L}^d\left(B_R \cap \{|X_1-X_2|>\gamma\}\right) \lesssim{}& \frac{1}{\delta\log\left(1+\frac{\gamma}{\delta}\right)}\|b_1-b_2\|_{L^1_t(L^1(B_{R+T\|b_2\|_\infty}))}\\ &+\frac{C_\varepsilon+\varepsilon\left(1+\log\left(\frac{C\|b_1\|_{L^\infty}}{\delta\varepsilon}\right)\right)}{\log\left(1+\frac{\gamma}{\delta}\right)}.\end{aligned}$$

By choosing first ε and then δ, one can prove existence, uniqueness, stability, and compactness as in the $p > 1$ case.

4. Sharp regularity of the solutions

In this final section, we will determine the sharp regularity of the solution to the continuity equation in our setting by deriving it from a suitable regularity property of the Regular Lagrangian Flow. This result was established by Brue and Nguyen in [**BN21**]. We will measure the regularity of the solution via a logarithmic version of the Gagliardo seminorm, which defines a scale of log-Sobolev functionals. For $p \ge 0$, the functional is given by

$$\left(\int_{\mathbb{R}^d}\int_{B_{1/3}} \frac{|f(x+h)-f(x)|^2}{|h|^d\log(1/|h|)^{1-p}}\,dh\,dx\right)^{1/2}. \tag{52}$$

This is inspired by the classical Gagliardo seminorm used in the theory of fractional Sobolev spaces, defined for $s \in (0,1)$ by

$$\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{|f(x+h)-f(x)|^2}{|h|^{d+2s}}\,dh\,dx\right)^{1/2}, \tag{53}$$

which provides a measure of the L^2-norm of the derivative of order s of the function f. We will prove the positive part of following theorem.

Theorem 3.1 Let $p > 1$ be fixed, and let $b \in L^1([0,T]; W^{1,p}(\mathbb{R}^d;\mathbb{R}^d))$ be a bounded divergence-free vector field. Then, for every initial datum $u_0 \in L^\infty(\mathbb{R}^d) \cap W^{1,1}(\mathbb{R}^d)$ with $\|u_0\|_{L^\infty} \le 1$, the solution $u \in L^\infty([0,T]\times\mathbb{R}^d)$ to the continuity equation (23) satisfies

$$\int_{B_{1/3}}\int_{\mathbb{R}^d} \frac{|u_t(x+h)-u_t(x)|^2}{|h|^d\log(1/|h|)^{1-p}}\,dx\,dh \le \int_0^t \|\nabla b_s\|_{L^p}^p\,ds + \|u_0\|_{W^{1,1}}^p + \|u_0\|_{L^1}.$$

Moreover, there exists a divergence-free vector field $b \in L^\infty([0,+\infty); W^{1,p}(\mathbb{R}^d))$ and an initial datum $u_0 \in L^\infty(\mathbb{R}^d) \cap W^{1,d}(\mathbb{R}^d)$, such that the solution $u \in L^\infty([0,+\infty) \times \mathbb{R}^d)$ of the Cauchy problem (23) satisfies

$$\int_{B_{1/2}} \int_{\mathbb{R}^d} \frac{|u_t(x+h) - u_t(x)|^2}{|h|^d \log(1/|h|)^\gamma} \, dx \, dh = \infty, \quad \text{for any } t > 0, \quad \text{for any } \gamma < 1 - p.$$

We start with a proposition concerning the (unique) Regular Lagrangian Flow.

Proposition 3.2 Under the assumptions of Theorem 3.1, there exists a measurable function $g(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R} \cup +\infty$ such that

$$e^{-g_t(x) - g_t(y)} \le \frac{|X_t(x) - X_t(y)|}{|x-y|} \le e^{g_t(x) + g_t(y)} \quad \text{for any } x, y \in \mathbb{R}^d \text{ and } t \in [0,T] \tag{54}$$

and

$$\|g_t\|_{L^p} \lesssim_{p,d} \int_0^t \|\nabla b_s\|_{L^p} \, ds \quad \forall t \in [0,T]. \tag{55}$$

Proof. We define all functions pointwise, as in the proof of Proposition A.29. Let $\varepsilon > 0$; then, for almost every $x, y \in \mathbb{R}^d$, we consider the quantity:

$$\left| \log \left(\frac{\varepsilon + |X_t(x) - X_t(y)|}{\varepsilon + |x-y|} \right) \right| = \left| \int_0^t \frac{d}{ds} \log\left(\varepsilon + |X_s(x) - X_s(y)|\right) \, ds \right|.$$

Now, using the chain rule and the fact that

$$\frac{d}{ds} |X_s(x) - X_s(y)| = \frac{X_s(x) - X_s(y)}{|X_s(x) - X_s(y)|} \cdot \left(\dot{X}_s(x) - \dot{X}_s(y) \right),$$

we obtain the estimate:

$$\frac{d}{ds} \log\left(\varepsilon + |X_s(x) - X_s(y)|\right) \le \frac{|b_s(X_s(x)) - b_s(X_s(y))|}{\varepsilon + |X_s(x) - X_s(y)|}.$$

Hence,

$$\left| \log \left(\frac{\varepsilon + |X_t(x) - X_t(y)|}{\varepsilon + |x-y|} \right) \right| \le \int_0^t \frac{|b_s(X_s(x)) - b_s(X_s(y))|}{|X_s(x) - X_s(y)|} \, ds.$$

By using inequality (43) and letting $\varepsilon \to 0$, we obtain

$$\left| \log \left(\frac{|X_t(x) - X_t(y)|}{|x-y|} \right) \right| \le C_d \int_0^t \left[M(|\nabla b_s|)(X_s(x)) + M(|\nabla b_s|)(X_s(y)) \right] ds.$$

Define the function $g_t : \mathbb{R}^d \to [0, +\infty]$ by

$$g_t(x) := C_d \int_0^t M(|\nabla b_s|)(X_s(x)) \, ds \quad \text{for } x \in \mathbb{R}^d \setminus N,$$

and set $g_t(x) := +\infty$ for $x \in N$, where N is a set of measure zero on which the Regular Lagrangian Flow does not solve the ODE. Finally, the integrability condition on g, (55), follows easily from the boundedness of the maximal function in L^p, together with the Minkowski integral inequality. □

Note that (54) shows that, if we restrict ourselves to a sub-level set of g_t, then the flow is Lipschitz. In this sense, the previous proposition implies that, by removing a set of suitably small measure, the flow becomes Lipschitz on its complement. This property is known as the *Lusin–Lipschitz* property, in analogy with Lusin's Theorem.

Thanks to the previous proposition we can deduce something similar for the solution of the continuity equation

Proposition 3.3 Let $b \in L^1([0,T]; W^{1,p}(\mathbb{R}^d;\mathbb{R}^d))$ be a bounded, divergence-free vector field with $p > 1$. Then, there exists a measurable function $\tilde{g}_t(x) = \tilde{g}(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ such that, for every initial datum $\bar{u} \in W^{1,1}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, there exists a representative $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ of the (unique) solution in $L^\infty([0,T] \times \mathbb{R}^d)$ to the continuity equation (CE), satisfying

$$|u_t(x) - u_t(y)| \leq |x-y| e^{\tilde{g}_t(x) + \tilde{g}_t(y)}, \quad \forall x, y \in \mathbb{R}^d, \ \forall t \in [0,T], \tag{56}$$

and

$$\|\tilde{g}_t\|_{L^p(\mathbb{R}^d)} \leq C_{p,d} \left(\int_0^t \|\nabla b_s\|_{L^p(\mathbb{R}^d)} \, ds + \|\bar{u}\|_{W^{1,1}(\mathbb{R}^d)} \right), \quad \forall t \in [0,T],$$

for a constant $C_{p,d} > 0$ depending only on p and d.

Proof. Let X be the (unique) Regular Lagrangian Flow associated to the vector field b. It can be proved (see Theorem 6.2 in [**Amb04**]) that the map $x \mapsto X(t,x)$ is essentially invertible, i.e., there exists a measurable map

$$Y : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$$

such that

$$X(t, Y(t,x)) = Y(t, X(t,x)) = x \quad \text{for } \mathscr{L}^d\text{-a.e. } x \in \mathbb{R}^d. \tag{57}$$

It is easy to see that Y_t satisfies (54) with $\bar{g}_t(x) := g_t(Y_t(x))$ when x satisfies (57), and $\bar{g}_t(x) := +\infty$ otherwise. By using that also Y_t preserves the Lebesgue measure we have from (55) that

$$\|\bar{g}_t\|_{L^p(\mathbb{R}^d)} \lesssim_{p,d} \int_0^t \|\nabla b_s\|_{L^p(\mathbb{R}^d)} \, ds, \quad \forall t \in [0,T].$$

We now use the fact that $u_t = \bar{u}(Y_t)$ for almost every $x \in \mathbb{R}^d$ (this corresponds to (22) in the low regularity setting), together with (43) and (54) for Y_t, to obtain the following estimate:

$$\begin{aligned} \frac{|u_t(x) - u_t(y)|}{|x-y|} &= \frac{|\bar{u}(Y_t(x)) - \bar{u}(Y_t(y))|}{|Y_t(x) - Y_t(y)|} \cdot \frac{|Y_t(x) - Y_t(y)|}{|x-y|} \\ &\leq C_d \left[M(|\nabla \bar{u}|)(Y_t(x)) + M(|\nabla \bar{u}|)(Y_t(y)) \right] \cdot \frac{|Y_t(x) - Y_t(y)|}{|x-y|} \\ &\leq C_d \left[M(|\nabla \bar{u}|)(Y_t(x)) + M(|\nabla \bar{u}|)(Y_t(y)) \right] e^{\bar{g}_t(x) + \bar{g}_t(y)}. \end{aligned}$$

Moreover notice that by defining $\tilde{g}_t(x) = 2\bar{g}_t(x) + 2\log\left(\max\{C_d M(|\nabla \bar{u}|)(Y_t(x)), 1\}\right)$. we have that, for $x, y \in \mathbb{R}^d$ and $t \in [0,T]$, one has

$$C_d \left(M|\nabla \bar{u}|(Y_t(x)) + M|\nabla \bar{u}|(Y_t(y)) \right) e^{\bar{g}_t(x) + \bar{g}_t(y)} \leq e^{\tilde{g}_t(x) + \tilde{g}_t(y)}.$$

By combining it with the previous inequality, we have exactly

$$|u_t(x) - u_t(y)| \leq |x-y| e^{\tilde{g}_t(x) + \tilde{g}_t(y)}, \quad \forall x, y \in \mathbb{R}^d, \ \forall t \in [0,T],$$

as wanted. Finally, the inequality for $\|\tilde{g}\|_{L^p}$ follows directly from the corresponding bound for $\bar{g}$ and the definition of $\tilde{g}$. $\square$

Thanks to the previous proposition, we are led to study the regularity of functions f satisfying

$$|f(x) - f(y)| \leq |x-y| e^{g(x) + g(y)}$$

for some $g \in L^p$. The following proposition provides a characterization of their regularity.

Recall that $a \wedge b = \min(a,b)$.

Proposition 3.4 Let $p \geq 1$ be fixed. For any $f \in L^1(\mathbb{R}^d)$ satisfying the exponential Lusin–Lipschitz estimate

$$|f(x) - f(y)| \leq |x-y| \exp\{g(x) + g(y)\} \quad \forall x, y \in \mathbb{R}^d, \tag{58}$$

for some $g \in L^p(\mathbb{R}^d)$, it holds that

$$\int_{B_{1/3}} \int_{\mathbb{R}^d} \frac{1 \wedge |f(x+h) - f(x)|^2}{|h|^d \log(1/|h|)^{1-p}} \, dx \, dh \lesssim_{p,d} \|g\|_{L^p}^p + \|f\|_{L^1}. \tag{59}$$

PROOF. We will first prove that for every $h \in \mathbb{R}^d$ with $|h| \leq 1/e$, it holds that

$$\int_{\mathbb{R}^d} 1 \wedge |f(x+h) - f(x)|^2 \, dx \lesssim_d |h|^2 \int_1^{\log(1/|h|)} e^{2\lambda} |\{2g > \lambda\}| \, d\lambda + |h| \|f\|_{L^1}. \tag{60}$$

By using the layer cake formula and (58), we have

$$\begin{aligned} \int_{\mathbb{R}^d} 1 \wedge |f(x+h) - f(x)|^2 \, dx &= 2 \int_0^{e|h|} t \, |\{x : |f(x+h) - f(x)| > t\}| \, dt \\ &\quad + 2 \int_{e|h|}^1 t \, |\{x : |f(x+h) - f(x)| > t\}| \, dt \\ &\lesssim |h| \|f\|_{L^1} + \int_{e|h|}^1 t \, |\{x : |f(x+h) - f(x)| > t\}| \, dt \\ &\quad + \int_{e|h|}^1 t \, |\{x : g(x) + g(x+h) > \log\left(\frac{t}{|h|}\right)\}| \, dt. \end{aligned}$$

Now, estimating

$$\int_{e|h|}^1 t \, |\{x : g(x) + g(x+h) > \log\left(\frac{t}{|h|}\right)\}| \, dt \leq 2 \int_{e|h|}^1 t \, |\{x : 2g(x) > \log\left(\frac{t}{|h|}\right)\}| \, dt.$$

Finally, setting $\lambda = \log\left(\frac{t}{|h|}\right)$ and changing variables in the integral above, we conclude the proof of (60). We now use this inequality to prove the proposition, to simplify the notation we use $\mu(\lambda) := |\{2g > \lambda\}| \, d\lambda$. By using (60) we have

$$\begin{aligned} &\int_{B_{1/e}} \int_{\mathbb{R}^d} \frac{1 \wedge |f(x+h) - f(x)|^2}{|h|^d \log(1/|h|)^{1-p}} \, dx \, dh \\ &\lesssim \int_{B_{1/e}} \frac{\log(1/|h|)^{p-1}}{|h|^d} \int_1^{\log(1/|h|)} e^{2\lambda} \, d\mu(\lambda) \, dh + \int_{B_{1/e}} |h| \|f\|_{L^1} \, dh. \end{aligned}$$

We first change variables in spherical coordinates with radius $|h| = r$ and then according to $log(1/r) = t$ to get

$$\begin{aligned} &\int_{B_{1/e}} \frac{\log(1/|h|)^{p-1}}{|h|^d} \int_1^{\log(1/|h|)} e^{2\lambda} \, d\mu(\lambda) \, dh \\ &\lesssim \int_0^{1/e} \int_{\mathbb{S}^{d-1}} \frac{\log(1/r)^{p-1}}{r^d} \int_1^{\log(1/r)} e^{2\lambda} \, d\mu(\lambda) \, r^{d-1} \, d\omega \, dr \\ &\lesssim \int_0^{1/e} \frac{\log(1/r)^{p-1}}{r} \int_1^{\log(1/r)} e^{2\lambda} \, d\mu(\lambda) \, dr \\ &\lesssim \int_1^\infty e^{-2t} t^{p-1} \left(\int_1^t e^{2\lambda} \, d\mu(\lambda) \right) dt. \end{aligned}$$

We now apply Tonelli theorem to get

$$\int_1^\infty e^{-2t} t^{p-1} \int_1^t e^{2\lambda} \, d\mu(\lambda) \, dt = \int_1^\infty e^{2\lambda} \int_\lambda^\infty e^{-2t} t^{p-1} \, dt \, d\mu(\lambda).$$

It is easy to see that by integration by parts (and $\lambda \geq 1$) that

$$e^{2\lambda} \int_\lambda^\infty e^{-2t} t^{p-1} \, dt \lesssim \lambda^{p-1}.$$

Thus by doing layer cake formula in the other way we have

$$\int_1^\infty e^{-2t} t^{p-1} \int_1^t e^{2\lambda} \, d\mu(\lambda) \, dt \lesssim \int_1^\infty \lambda^{p-1} d\mu(\lambda) \lesssim \|g\|_{L^p}^p.$$

Finally, by putting all together we have

$$\int_{B_{1/e}} \int_{\mathbb{R}^d} \frac{1 \wedge |f(x+h) - f(x)|^2}{|h|^d \log(1/|h|)^{1-p}} \, dx \, dh \lesssim \|g\|_{L^p}^p + \|f\|_{L^1}$$

which completed the proof. □

We conclude the proof of Theorem 3.1 by combining the results established in the previous propositions. Specifically, we apply Proposition 3.3 in conjunction with Proposition 3.4, and we observe that the L^1-norm of u_t remains constant over time, which is equal to that of $\bar{u}$. This property arises from the fact that bounded solutions of the continuity equation are renormalized.

CHAPTER 4

Osgood meets DiPerna-Lions

1. Introduction

In the first part of this chapter, we define what we mean by functions with slightly less than a derivative in L^p. We recall the following fact, which provides a characterization of Sobolev functions:

Proposition 4.1 Let $f\colon \mathbb{R}^d \to \mathbb{R}$ be a function in L^p. Then $f \in W^{1,p}(\mathbb{R}^d)$ for $p > 1$ if and only if there exists a function $g \in L^p(\mathbb{R}^d)$ such that for almost every pair $x, y \in \mathbb{R}^d$, we have

$$|f(x) - f(y)| \leq |x - y|(g(x) + g(y)). \tag{61}$$

Proof. The "only if" part follows easily from (A.29) and Theorem A.27. For the converse, we follow an argument of Hajłasz [**Ha96**]. By using the Riesz Representation Theorem, it suffices to show that there exists $h \in L^p(\mathbb{R}^d)$ such that for all $\varphi \in C_c^\infty(\mathbb{R}^d)$, we have

$$\left| \int_{\mathbb{R}^d} f \, \frac{\partial \varphi}{\partial x_1} \right| \leq \int_{\mathbb{R}^d} |h| \, \varphi.$$

We now fix $\varepsilon > 0$, and we integrate in mean both sides of (61) first with respect to y and then with respect to x, over the ball $B := B_z(\varepsilon)$. This yields

$$\fint_B \left| f - \fint_B f \right| \lesssim \varepsilon \fint_B g. \tag{62}$$

To conclude the proof, we use a standard approximation of identity argument. Let ψ be a smooth function supported in the unit ball and satisfying $\int_{\mathbb{R}^d} \psi = 1$. As usual, we define the rescaled mollifier

$$\psi_\varepsilon(x) := \frac{1}{\varepsilon^d} \psi\left(\frac{x}{\varepsilon}\right).$$

By standard properties of mollifiers, we have

$$\int_{\mathbb{R}^d} f \, \frac{\partial \varphi}{\partial x_1} = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} (f * \psi_\varepsilon) \, \frac{\partial \varphi}{\partial x_1} = - \lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} \varphi \, \left(\frac{\partial \psi_\varepsilon}{\partial x_1} * f \right).$$

Now observe that $\int_{\mathbb{R}^d} \frac{\partial \psi_\varepsilon}{\partial x_1} = 0$, so

$$\left(\frac{\partial \psi_\varepsilon}{\partial x_1} * f \right)(x) = \left(\frac{\partial \psi_\varepsilon}{\partial x_1} * \left(f - \fint_{B_x(\varepsilon)} f \right) \right)(x).$$

Using the estimate (62) we have

$$\left| \left(\frac{\partial \psi_\varepsilon}{\partial x_1} * f \right)(x) \right| \lesssim \int_{B_x(\varepsilon)} |f(y) - \fint_{B_x(\varepsilon)} f| \, \frac{1}{\varepsilon^{d+1}} \, dy \lesssim \fint_{B_x(\varepsilon)} g \leq Mg(x).$$

Since $Mg \in L^p(\mathbb{R}^d)$, it follows that

$$\left| \int_{\mathbb{R}^d} f \, \frac{\partial \varphi}{\partial x_1} \right| \lesssim \int_{\mathbb{R}^d} \varphi \, Mg.$$

This completes the proof. □

Inspired by the previous proposition it is natural to consider the following class of functions.

Definition 4.2. Let ω be an Osgood modulus of continuity and $p \geq 1$. We say that a function $f \in L^p(\mathbb{R}^d, \mathbb{R}^d)$ belongs to $W^{\omega,p}(\mathbb{R}^d, \mathbb{R}^d)$ if there exists a function $g \in L^p(\mathbb{R}^d)$ such that for almost every pair of points $x, y \in \mathbb{R}^d$, we have

$$|f(x) - f(y)| \leq \omega(|x-y|)(g(x) + g(y)). \tag{63}$$

Remark 4.3. Note that $\omega(x) = x$ is the Sobolev space case.

In a natural way we extend this definition to define vector field in $L^1_t(W^{\omega,p}_x)$.

Definition 4.4. We say that a vector field $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ belongs to $L^1_t(W^{\omega,p}_x)$ if for almost every $t \in [0,T]$, there exists a function $g_t \in L^p(\mathbb{R}^d)$ such that

$$|b(t,x) - b(t,y)| \leq \omega(|x-y|)\big(g_t(x) + g_t(y)\big) \quad \text{for a.e. } x, y \in \mathbb{R}^d,$$

and

$$\int_0^T \|g_t\|_{L^p}\, dt < +\infty.$$

There are several motivations for studying this case of vector fields. First, as in the classical setting, one would like to understand whether it is possible to improve upon a classical threshold that appears to be almost necessary — such as the local Lipschitz condition discussed in Chapter 1. Here, we aim to refine the breakthrough result of DiPerna and Lions on Sobolev vector fields. Another interesting motivation arises from the following computation. Let $f \in W^{1,p}$. One can prove, by mimicking the same proof of (43), that

$$|f(x) - f(y)| \leq \omega(|x-y|)\big(M_\omega(\nabla f)(x) + M_\omega(\nabla f)(y)\big)$$

where

$$M_\omega(g)(x) := \sup_{R>0} \frac{R}{\omega(R)} \fint_{B_R(x)} |g(y)|\, dy.$$

By using that ω is concave, and fixing K, one can perform an interpolation argument:

$$\sup_{K \geq R > 0} \frac{R}{\omega(R)} \fint_{B_R(x)} |g(y)|\, dy \leq \frac{K}{\omega(K)} M(g)(x).$$

In addition,

$$\sup_{R>K} \frac{R}{\omega(R)} \fint_{B_R(x)} |g(y)|\, dy \lesssim \frac{K}{\omega(K)} \frac{\|f\|_{L^1}}{K^d}.$$

Assuming $\|f\|_{L^1} = 1$ and optimizing in K, we obtain

$$M_\omega(g)(x) \lesssim \frac{[M(g)(x)]^{1-\frac{1}{d}}}{\omega\big([M(g)(x)]^{-\frac{1}{d}}\big)}.$$

In particular, if we define the (increasing) function

$$\varphi(t) := \frac{t^{1-\frac{1}{d}}}{\omega(t^{-\frac{1}{d}})} \leq t$$

and its inverse $\psi(t) \geq t$, we have

$$\psi\big(M_\omega(g)(x)\big) \lesssim M(g)(x).$$

Thus we are "giving away" a part of the derivative to gain better integrability. One motivation for introducing this type of vector field was to obtain better estimates in the borderline case $p = 1$. Unfortunately, this trade-off between derivative and integrability is not sufficient to compensate for the gap between $L^{1,\infty}$ and L^1.

2. Regular Lagrangian Flow

In this section, we generalize some of the results from Chapter 3 by proving that, for vector fields as described in the introduction to this chapter (i.e., in $L^1_t(W^{\omega,p}_x)$, bounded, and divergence-free), there exists a unique Regular Lagrangian Flow. The following proof is a natural extension of [**CL08**]. While working on this thesis, the author independently arrived at the result presented in this section, which was later discovered to have already been proved in [**LL15**].

To mimic the function $\log\left(1+\frac{x}{\delta}\right)$ used in the proof by Crippa and De Lellis, it is natural to introduce the following function for each Osgood modulus of continuity:

$$Q_\delta(x) := \int_0^x \frac{ds}{\omega(s)+\delta}. \tag{64}$$

In particular, when $\omega(s)=s$, this function coincides with $\log\left(1+\frac{x}{\delta}\right)$.

After this natural modification, the proof proceeds essentially as in the original case. In particular, we aim to estimate the following quantity, for a fixed $R>0$ and for X_1, X_2 regular Lagrangian flows associated to the vector fields b_1 and b_2 with b_1 satisfying the Osgood derivative condition:

$$\Phi_\delta(t) := \int_{B_R} Q_\delta\left(|X_1(t,x)-X_2(t,x)|\right)\,dx.$$

As before, we differentiate $\Phi_\delta(t)$ and proceed with similar estimates. However, in the step where we previously used the maximal function estimate, we now exploit the assumption that the vector field has a L^p modulus of continuity given by an Osgood function ω. This yields:

$$\begin{aligned}
\Phi'_\delta(t) &\le \int_{B_R} \frac{\left|\frac{d}{dt}X_1 - \frac{d}{dt}X_2\right|}{\delta+\omega(|X_1-X_2|)}\,dx \\
&= \int_{B_R} \frac{|b_1^t(X_1)-b_2^t(X_2)|}{\delta+\omega(|X_1-X_2|)}\,dx \\
&\le \int_{B_R} \frac{|b_1^t(X_1)-b_1^t(X_2)+b_1^t(X_2)-b_2^t(X_2)|}{\delta+\omega(|X_1-X_2|)}\,dx \\
&\le \int_{B_R} \frac{|b_1^t(X_1)-b_1^t(X_2)|}{\delta+\omega(|X_1-X_2|)}\,dx + \int_{B_R} \frac{|b_1^t(X_2)-b_2^t(X_2)|}{\delta+\omega(|X_1-X_2|)}\,dx \\
&=: (I)+(II).
\end{aligned}$$

We start from (I)

$$\begin{aligned}
\int_{B_R} \frac{|b_1^t(X_1)-b_1^t(X_2)|}{\delta+\omega(|X_1-X_2|)}\,dx &\le \int_{B_R} \left(g_t(X_1)+g_t(X_2)\right)dx \\
&\le \int_{B_{R+T\|b_1\|_\infty}} g_t(y)\,dy + \int_{B_{R+T\|b_2\|_\infty}} g_t(y)\,dy \\
&\le C_1\|g_t\|_{L^p},
\end{aligned}$$

where the constant C_1 depends on $R, \|b_1\|_\infty, \|b_2\|_\infty, d, T$.

We have used the fact that the flow preserves the Lebesgue measure, Hölder's inequality and the property imposed on b_1.

Now we estimate the other term:

$$(II) \le \int_{B_R} \frac{|b_1^t(X_2)-b_2^t(X_2)|}{\delta}\,dx \le \frac{1}{\delta}\|b_1^t-b_2^t\|_{L^1(B_{R+T\|b_2\|_\infty})}.$$

Putting everything together and integrating in time (and noticing that $\Phi_\delta(0)=0$), we obtain:

$$\Phi_\delta(t) \le \frac{1}{\delta}\|b_1-b_2\|_{L^1_t(L^1(B_{R+T\|b_2\|_\infty}))} + C_1\|g\|_{L^1_t(L^p_x)}. \tag{65}$$

At this point we use the same idea as the other time i.e (44):

$$\Phi_\delta(t) \geq \int_{B_R \cap \{|X_1 - X_2| \geq \gamma\}} Q_\delta(\gamma)\, dx = Q_\delta(\gamma)\, \mathscr{L}^d \left(B_R \cap \{|X_1 - X_2| \geq \gamma\}\right). \tag{66}$$

By combining (65) and (66), the proofs of existence, uniqueness, and compactness follow exactly as in Chapter 3, with the additional use of the Osgood condition, namely

$$\lim_{\delta \to 0} Q_\delta(x) = +\infty \quad \text{for each } x > 0.$$

Note that, after we have proved existence and uniqueness for the Regular Lagrangian Flows, Theorem 2.17 guarantees that positive solutions of the continuity equation are unique within the class of positive solutions.

3. Renormalized solution

In this section, we prove that, under a suitable regime induced by the Osgood modulus of continuity ω, bounded solutions are renormalized as in the classical theory of DiPerna and Lions. This is joint work with Guido De Philippis. The scheme of the proof is as follows: using the Littlewood–Paley decomposition, we show that solutions transported by a vector field whose "derivative" satisfies an Osgood condition possess a certain regularity. We then exploit this regularity to control the commutator, proving that it vanishes in the limit in the L^1_{loc} sense. This allows us to apply the DiPerna–Lions framework and conclude that solutions of the continuity are renormalized even in this setting of vector fields. For clarity, we work on the periodic torus; however, the proof can be adapted to work locally in $\mathbb{R}^d$.

3.1. Littlewood-Paley Decomposition. In this subsection, we introduce the Littlewood–Paley decomposition, a useful way to decompose a function into pieces that are well localized in the frequency space. For further details and proofs one can see the Appendix. An easy example that illustrates the type of result we aim to obtain is the following. Let f_k be a sequence of functions indexed by $k \in \mathbb{Z}_+$, each with Fourier support contained in the annulus $2^k < |\xi| < 2^{k+1}$. Then, by Parseval's formula and the fact that the Fourier supports of the f_k are disjoint, we obtain

$$\left\| \sum_k f_k \right\|_{L^2}^2 = \sum_k \|f_k\|_{L^2}^2.$$

Unfortunately, one can see that for $p \neq 2$, equalities of this type no longer hold. This questions are related to a central theme in modern harmonic analysis that is the study of inequalities of the form

$$\left\| \sum_k f_k \right\|_{L^p} \lesssim \left(\sum_k \|f_k\|_{L^p}^2 \right)^{1/2}, \quad 1 < p < \infty,$$

under suitable assumptions on the functions f_k. These are commonly referred to as *decoupling estimates*. For a more broad overview of this topic, see [**Gut22**]. The Fourier transform of a function θ on $\mathbb{T}^d = \mathbb{R}^d/(\mathbb{Z}^d)$ and its Fourier representation are defined as

$$\widehat{\theta}(\eta) = \int_{\mathbb{T}^d} e^{-2\pi i \eta \cdot x} \theta(x)\, dx, \quad \theta(x) = \sum_{\eta \in \mathbb{Z}^d} e^{2\pi i \eta \cdot x} \widehat{\theta}(\eta),$$

for any wave number $\eta \in \mathbb{Z}^d$ and any $x \in \mathbb{T}^d$.

The Fourier transform of a function φ on $\mathbb{R}^d$ and its Fourier representation are defined as

$$\widehat{\varphi}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot x} \varphi(x)\, dx, \quad \varphi(x) = \int_{\mathbb{R}^d} e^{2\pi i \xi \cdot x} \widehat{\varphi}(\xi)\, d\xi,$$

for any frequency $\xi \in \mathbb{R}^d$ and any $x \in \mathbb{R}^d$.

Convolutions will always be understood on $\mathbb{R}^d$, which occasionally makes it necessary to extend functions periodically from $\mathbb{T}^d$ to $\mathbb{R}^d$. The convolution of θ and φ, as defined above, is the function

$$(\varphi * \theta)(x) = \int_{\mathbb{R}^d} \varphi(x-y)\theta(y)\,dy = \int_{\mathbb{T}^d} \left(\sum_{\eta\in\mathbb{Z}^d} \varphi(x-z-\eta)\right)\theta(z)\,dz, \tag{67}$$

which is periodic in x, and is thus interpreted as a function on $\mathbb{T}^d$. Its Fourier transform then satisfies the product rule:

$$\widehat{(\varphi * \theta)}(\eta) = \widehat{\varphi}(\eta)\,\widehat{\theta}(\eta), \tag{68}$$

for any $\eta \in \mathbb{Z}^d$.

Similarly, the product of two functions transforms into a discrete convolution when passing to the Fourier transform:

$$\widehat{(\theta\phi)}(\eta) = \sum_{\zeta\in\mathbb{Z}^d} \widehat{\theta}(\zeta)\,\widehat{\phi}(\eta-\zeta), \quad \text{for each } \eta \in \mathbb{Z}^d. \tag{69}$$

Moreover, we have the Parseval's identity:

$$\int_{\mathbb{T}^d} \theta(x)\overline{\phi(x)}\,dx = \sum_{\eta\in\mathbb{Z}^d} \widehat{\theta}(\eta)\overline{\widehat{\phi}(\eta)}. \tag{70}$$

We now define the Littlewood–Paley decomposition by first considering a radial Schwartz function φ whose Fourier transform $\widehat{\varphi}$ is supported in the unit ball $B_1(0)$. Moreover, $\widehat{\varphi}$ satisfies

$$0 \le \widehat{\varphi}(\xi) \le 1 \quad \text{for all } \xi,$$

and equals 1 on the closed ball $\overline{B_{\frac{1}{2}}(0)}$. This function will be used to construct a dyadic partition of unity in frequency space.

For each $k \in \mathbb{Z}$, we define

$$\varphi_k(x) := 2^{kd}\varphi(2^k x) - 2^{(k-1)d}\varphi(2^{k-1}x), \tag{71}$$

so that the Fourier transform $\widehat{\varphi}_k$ is supported in a dyadic annulus:

$$\operatorname{spt}\widehat{\varphi}_k \subset B_{2^k}(0) \setminus \overline{B_{2^{k-2}}(0)}. \tag{72}$$

As a result, each φ_k overlaps in frequency only with its immediate neighbors. Notice that (71) is defined in such a way that

$$\widehat{\varphi}_k(\xi) = \widehat{\varphi}(\frac{\xi}{2^k}) - \widehat{\varphi}(\frac{\xi}{2^{k-1}}). \tag{73}$$

From this last equation and the definition of φ we deduce that

$$\widehat{\varphi}_k(\xi) = \widehat{\varphi}_k(\xi)(\widehat{\varphi}_{k-1}(\xi) + \widehat{\varphi}_k(\xi) + \widehat{\varphi}_{k+1}(\xi))$$

since $\widehat{\varphi}_{k-1}(\xi) + \widehat{\varphi}_k(\xi) + \widehat{\varphi}_{k+1}(\xi) = 1$ on the support of $\widehat{\varphi}_k(\xi)$. By applying the inverse Fourier transform, we obtain

$$\varphi_k = \varphi_k * (\varphi_{k-1} + \varphi_k + \varphi_{k+1}). \tag{74}$$

Notice also that the family $\{\varphi_k\}_{k\in\mathbb{Z}}$ is constructed to form a partition of unity in frequency space, i.e.,

$$\sum_{k\in\mathbb{Z}} \widehat{\varphi}_k(\xi) = 1 \quad \text{for all } \xi \in \mathbb{R}^d \setminus \{0\}.$$

From now on, every function u on the torus will be assumed to be mean-free; that is,

$$\int_{\mathbb{T}^d} u(x)\,dx = 0.$$

For a mean-free function on the torus, we define the k-th Littlewood–Paley block f_k as

$$f_k := f * \varphi_k.$$

In addition to the Littlewood–Paley blocks, we consider the low-pass filters

$$\psi_k(x) := \varphi(x) + \sum_{j=1}^{k} \varphi_j(x),$$

and define the high-frequency parts as

$$f_k^{\geq} := f - f * \psi_{k-1}.$$

Similarly, we define

$$f_k^{\leq} := f * \psi_k,$$

so that

$$f = f_k^{\leq} + f_{k+1}^{\geq}.$$

Moreover, one can see that for $1 < p < +\infty$, if f is mean-free and in L^p, then

$$f = \sum_{k=1}^{\infty} f_k \quad \text{in } L^p(\mathbb{T}^d)$$

and

$$f_m^{\geq} = \sum_{k=m}^{\infty} f_k \quad \text{in } L^p(\mathbb{T}^d).$$

We will make use of this property several times in the context of regularity estimates.

One of the most important properties of the Littlewood Paley decomposition is that, for all integrability exponents $q \in (1, \infty)$,

$$\left\| \left(\sum_k u_k^2 \right)^{1/2} \right\|_{L^q} \sim \|u\|_{L^q}. \tag{75}$$

For a proof, see Theorem 6.1.2 in [**Gra14**].

Later, we will use only the upper bound, which holds for more general L^1-dilations $\widetilde{\varphi}_k = 2^{kd}\widetilde{\varphi}_0(2^k\cdot)$, provided that $\widetilde{\varphi}_0$ is a mean-zero Schwartz function:

$$\left\| \left(\sum_k (u * \widetilde{\varphi}_k)^2 \right)^{1/2} \right\|_{L^q} \lesssim \|u\|_{L^q}. \tag{76}$$

For a proof, see Theorem 6.1.2 in [**Gra14**].
The next lemma will be very useful several times.

Lemma 4.5 Let K be a Schwartz function, then for every $N > 0$ we have that

$$N^d \int_{\mathbb{R}^d} K(N(x-y)) f(y) dy \lesssim_K Mf(x) \tag{77}$$

Proof. Without loss of generality we can assume $K, f \geq 0$. We decompose the integral in several annuli defined by a dyadic partition.

$$\begin{aligned} N^d \int_{\mathbb{R}^d} K(N(x-y)) f(y)\, dy = N^d &\int_{N|x-y|\leq 1} K(N(x-y)) f(y)\, dy \\ &+ \sum_{k=0}^{\infty} N^d \int_{2^k \leq N|x-y| \leq 2^{k+1}} K(N(x-y)) f(y)\, dy \end{aligned}$$

We now use the definition of a Schwartz function, i.e., there exists $C > 0$ such that

$$K(x) \leq \frac{C}{1 + |x|^{d+1}} \quad \text{for all } x \in \mathbb{R}^d.$$

In particular, we have the following estimate:

$$N^d \int_{N|x-y|\leq 1} K(N(x-y)) f(y)\, dy \leq C N^d \int_{N|x-y|\leq 1} f(y)\, dy \lesssim Mf(x).$$

On the other hand, for the other integrals we have

$$\begin{aligned}
&N^d \int_{2^k \le N|x-y| \le 2^{k+1}} K(N(x-y))f(y)\,dy \\
&\le CN^d \int_{2^k \le N|x-y| \le 2^{k+1}} f(y)\frac{dy}{N^{d+1}|x-y|^{d+1}} \\
&\lesssim N^d \int_{2^k \le N|x-y| \le 2^{k+1}} f(y)\frac{dy}{2^{(k+1)(d+1)}} \\
&\lesssim N^d \int_{N|x-y| \le 2^{k+1}} f(y)\frac{dy}{2^{(k+1)(d+1)}} \\
&\lesssim \frac{Mf(x)}{2^{k+1}}.
\end{aligned}$$

Finally, by summing over k, we get that

$$N^d \int_{\mathbb{R}^d} K(N(x-y))f(y)\,dy \lesssim_K Mf(x) + Mf(x) \lesssim_K Mf(x),$$

which concludes the proof of the lemma. □

Remark 4.6. Note that applying this to $K = \varphi_k, \psi_k$ yields

$$|f_k(x)| + |f_k^{\le}(x)| \lesssim Mf(x), \tag{78}$$

where the implicit constant is independent of k and f.

In addition to the square function relation (75), there are similar estimates that relate the norm of derivatives, known as the *Bernstein inequalities*.

Proposition 4.7 For $1 \le p \le \infty$ and $s \in \mathbb{R}$, we have

$$\||\nabla|^s f_k\|_{L^p} \sim 2^{ks}\|f_k\|_{L^p}.$$

Moreover, for $s > 0$, the following inequalities hold:

$$\||\nabla|^s f_k^{\le}\|_{L^p} \lesssim 2^{ks}\|f\|_{L^p},$$

$$\|f_k^{\le}\|_{L^p} \lesssim 2^{-ks}\||\nabla|^s f\|_{L^p}.$$

Finally, for every $s \in \mathbb{R}$ and $1 < p < \infty$, the following equivalences hold:

$$\||\nabla|^s f\|_{L^p} \sim_{s,p} \left\| \left(\sum_k 2^{2ks}|f_k(x)|^2 \right)^{\frac{1}{2}} \right\|_{L^p},$$

and for $s > 0$

$$\||\nabla|^s f\|_{L^p} \sim_{s,p} \left\| \left(\sum_k 2^{2ks}|f_k^{\ge}(x)|^2 \right)^{\frac{1}{2}} \right\|_{L^p}.$$

3.2. Regularity: Preparatory Propositions. In this subsection, we prove that solutions to the continuity equation exhibit a certain regularity. This argument is inspired by [**MS24**], where the authors employed the Littlewood–Paley decomposition to re-establish the sharp regularity results for the continuity equation originally proved in [**BN21**]. Before defining the notion of regularity that we aim to establish, we introduce the function G, analogous to the function W defined in (9):

$$G(k) := W(2^{-k}) = \int_{2^{-k}}^1 \frac{ds}{\omega(s)}.$$

Note that $G(k) = k$ when $\omega(s) = s$. We extend G to real values by setting

$$G(x) := \int_{2^{-x}}^1 \frac{ds}{\omega(s)} \quad \text{for } x > 0,$$

and defining $G(x) := 0$ for $x \leq 0$. Given a function u, we define the norm $\|u\|_{B^{\omega,a}}$ by

$$\|u\|_{B^{\omega,a}} := \left(\sum_{k=1}^{\infty} G(k)^{2a}\|u_k\|_{L^2}^2\right)^{1/2}, \tag{79}$$

where u_k denotes the k-th Littlewood–Paley block of u.

Remark 4.8. Note that when $\omega(s) = s$, the norm reduces to

$$\|u\|_{B^{\log,a}} := \left(\sum_{k=1}^{\infty} k^{2a}\|u_k\|_{L^2}^2\right)^{1/2},$$

which is the norm used to characterize sharp regularity in the Sobolev case, as established in the two previously cited papers.

Before proving some preparatory propositions for the regularity estimates, we first establish some properties of $G(k)$. Recall that our Osgood modulus of continuity ω is continuous, increasing, and concave. The Osgood condition implies that

$$\lim_{k\to\infty} G(k) = +\infty.$$

Lemma 4.9 For each positive integer k, we have

$$G(k) \leq G(k+1) \leq G(k) + G(1).$$

Proof. The first inequality follows immediately since ω is positive on $(0,1)$. For the second inequality, we use the concavity of ω, which implies that for any $y < 1$,

$$\omega(x)y \leq \omega(xy).$$

We then compute:

$$G(k+1) = \int_{2^{-(k+1)}}^{1} \frac{ds}{\omega(s)} = \int_{2^{-k}}^{1} \frac{ds}{\omega(s)} + \int_{2^{-(k+1)}}^{2^{-k}} \frac{ds}{\omega(s)} = G(k) + \int_{2^{-(k+1)}}^{2^{-k}} \frac{ds}{\omega(s)}.$$

Using the change of variables $s = \frac{x}{2^k}$, the last integral becomes

$$\int_{\frac{1}{2}}^{1} \frac{2^{-k}dx}{\omega\left(\frac{x}{2^k}\right)} = \int_{\frac{1}{2}}^{1} \frac{dx}{2^k\omega\left(\frac{x}{2^k}\right)}.$$

By the concavity property above, this is bounded by

$$\int_{\frac{1}{2}}^{1} \frac{dx}{\omega(x)} = G(1).$$

Putting it all together, we get

$$G(k+1) \leq G(k) + G(1),$$

which completes the proof. □

This lemma shows that the function $G(k)$ tends to infinity as $k \to \infty$, but its growth is at most sublinear. Moreover, it is straightforward to see that $G(k)$ and $G(k+1)$ are comparable, in the sense that their ratio converges to 1 as $k \to \infty$. Next, we require another lemma, which is in the spirit of estimating an integral by the corresponding series.

Lemma 4.10 Let $a \geq 1$ and ω be as before. Then, for all integers $k \geq 1$, we have

$$G(k)^{2a} \lesssim \sum_{j=1}^{k} G(j)^{2a-1}\frac{2^{-j}}{\omega(2^{-j})} \lesssim \sum_{j=1}^{k+1} G(j)^{2a-1}\frac{2^{-j}}{\omega(2^{-j})} \lesssim G(k)^{2a}, \tag{80}$$

where the implicit constants depend only on ω and a, but not on k.

Proof. First, notice that, by what was said earlier, the two sums in the middle are comparable. The idea is that, after a change of variables, we will use the fact that—under suitable assumptions (which are satisfied in our case, essentially by what was said earlier) the sum and the integral of a function f are comparable:

$$\sum_{j=1}^{k} f(j) \sim \int_0^k f(x)\,dx.$$

We now compute the derivative of $G(x)^{2a}$ with respect to the variable x. One can see that

$$\frac{d}{dx}\left(G(x)^{2a}\right) = 2a\,G(x)^{2a-1}G'(x) = 2a\ln 2\,G(x)^{2a-1}\frac{2^{-x}}{\omega(2^{-x})}.$$

We choose this expression as our function $f(x)$ in the comparison between the sum and the integral, and the result follows. □

Sometimes we will not work with exactly the norm defined in (79), instead we will work with the equivalent norm

$$\|u\|_{\mathbb{B}^{\omega,a}} := \left(\sum_{k=1}^{\infty} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\left\|u_k^{\geq}\right\|_{L^2}^2\right)^{1/2} \tag{81}$$

We now prove this two norms are equivalent

Proposition 4.11 The norms $\|u\|_{B^{\omega,a}}$ and $\|u\|_{\mathbb{B}^{\omega,a}}$ are equivalent.

Proof. In order to prove the proposition, we need some inequalities between the Littlewood–Paley blocks. By (72), the Littlewood–Paley terms are orthogonal in L^2 unless the indices differ by less than one. We can therefore write

$$\left\|\sum_{k\geq j} u_k\right\|_{L^2}^2 = \sum_{k\geq j}\sum_{\ell=k-1}^{k+1}\int_{\mathbb{T}^d} u_k u_\ell\,dx \lesssim \sum_{k\geq j-1}\|u_k\|_{L^2}^2.$$

where in the last line we have used that $ab \lesssim a^2+b^2$.
Let now $k \geq j$, again by (72), we have that

$$\|u_k\|_{L^2}^2 = \left\|\sum_{m\geq j-1} u_k * \varphi_m\right\|_{L^2}^2,$$

and summing over k gives

$$\sum_{k\geq j}\|u_k\|_{L^2}^2 = \sum_{k\geq j}\left\|\sum_{m\geq j-1} u_k*\varphi_m\right\|_{L^2}^2 \leq \sum_{k\geq 1}\left\|\left(\sum_{m\geq j-1} u_m\right)*\varphi_k\right\|_{L^2}^2 \lesssim \left\|\sum_{m\geq j-1} u_m\right\|_{L^2}^2,$$

where in the last inequality (75) with $q = 2$ was used. We are now ready to conclude the proof by using (80) together with the previous inequality concerning the Littlewood–Paley pieces. Note that we also use the fact that the function is mean-free. In particular we have that

$$\begin{aligned}
\|u\|_{\mathbb{B}^{\omega,a}}^2 &= \sum_{k=1}^{\infty} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\|u_k^{\geq}\|_{L^2}^2 \\
&\lesssim \sum_{k=1}^{\infty} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\sum_{m\geq k-1}\|u_m\|_{L^2}^2 \\
&= \sum_{m=1}^{\infty}\left(\sum_{k=1}^{m+1} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\right)\|u_m\|_{L^2}^2 \\
&\lesssim \sum_{m=1}^{\infty} G(m)^{2a}\|u_m\|_{L^2}^2 = \|u\|_{B^{\omega,a}}^2.
\end{aligned}$$

The reverse inequality follows by a similar argument. □

We conclude this subsection with some interpolation inequalities that will be used throughout the main proof of the regularity estimate.

Proposition 4.12 Let $b, a \geq 0$ and $r \geq 2$ be given, and let $u \in L^\infty(\mathbb{T}^d) \cap B^{\omega,a}$. Let $\eta_k = 2^{kd}\eta(2^k\cdot)$ be a family of mean-free Schwartz functions for which $\widehat{\eta_1}$ is compactly supported.

a) If $2a = br$, then it holds

$$\|\sup_{k\geq 0} G(k)^b |u_k^{\geq}|\|_{L^r} \lesssim \|u\|_{L^\infty}^{1-\frac{b}{a}} \|u\|_{B^{\omega,a}}^{\frac{b}{a}}. \tag{82}$$

b) If $2a = br$, then it holds

$$\|\sup_{k\geq 0} G(k)^b |u_k|\|_{L^r} \lesssim \|u\|_{L^\infty}^{1-\frac{b}{a}} \|u\|_{B^{\omega,a}}^{\frac{b}{a}}. \tag{83}$$

c) If $b < \frac{2a}{r}$, then

$$\left\| \left(\sum_{k\geq 1} G(k)^{2b} |u * \eta_k|^2 \right)^{\frac{1}{2}} \right\|_{L^r} \lesssim_{\eta_1} \|u\|_{L^\infty}^{1-\frac{b}{a}} \|u\|_{B^{\omega,a}}^{\frac{b}{a}}. \tag{84}$$

Proof. We first prove (a). By (80), we obtain

$$G(k)^{br} \lesssim \sum_{j=1}^{k} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})},$$

and therefore

$$\begin{aligned}
\left\| \sup_{k\geq 0} G(k)^b |u_k^{\geq}| \right\|_{L^r}^r &= \left\| \sup_{k\geq 0} G(k)^{br} |u_k^{\geq}|^r \right\|_{L^1} \\
&\lesssim \left\| \sup_{k\geq 0} \sum_{j=1}^{k} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} |u_k^{\geq}|^r \right\|_{L^1} \\
&\leq \left\| \sum_{j\geq 1} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} \sup_{k\geq j} |u_k^{\geq}|^r \right\|_{L^1},
\end{aligned}$$

where in the last line we moved the supremum inside the sum.

We now proceed by moving the norm inside, using the triangle inequality in L^1:

$$\left\| \sum_{j=1}^{k} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} \sup_{k\geq j} |u_k^{\geq}|^r \right\|_{L^1} \leq \sum_{j\geq 1} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} \left\| \sup_{k\geq j} |u_k^{\geq}|^r \right\|_{L^1}.$$

We are now interested in estimating $\|\sup_{k\geq j} |u_k^{\geq}|\|_{L^r}$. To clarify the structure of the argument, we consider the sub-problem of estimating

$$\left\| \sup_{k\geq j} |f_k^{\geq}| \right\|_{L^r}$$

for a mean-free function $f \in L^r$.

In particular, we have that

$$\left\| \sup_{k\geq j} |f_k^{\geq}| \right\|_{L^r} = \left\| \sup_{k\geq j} |f - f_{k-1}^{\leq}| \right\|_{L^r} \leq \|f\|_{L^r} + \left\| \sup_{k\geq j-1} |f_k^{\leq}| \right\|_{L^r} \lesssim \|f\|_{L^r},$$

where we have used the pointwise estimate (78)

$$|f_k^{\leq}(x)| \leq Mf(x),$$

and the fact that

$$\|Mf\|_{L^r} \lesssim \|f\|_{L^r},$$

i.e the boundedness of the Hardy–Littlewood maximal operator on L^r.

By applying the previous estimate to $f = u_{j-2}^{\geq}$ and using the properties of the Littlewood Paley decomposition we can thus estimate

$$\left\| \sup_{k\geq j} G(k)^b |u_k^{\geq}| \right\|_{L^r}^r \lesssim \sum_{j\geq 1} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} \left\| u_{j-2}^{\geq} \right\|_{L^r}^r .$$

Finally, by applying the standard interpolation inequality for the L^r norm between L^2 and L^∞, we obtain:

$$\begin{aligned}\sum_{j\geq 1} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} \left\| u_{j-2}^{\geq} \right\|_{L^r}^r &\lesssim \sum_{j\geq 1} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} \left\| u_{j-2}^{\geq} \right\|_{L^2}^2 \left\| u_{j-2}^{\geq} \right\|_{L^\infty}^{r-2} \\ &\lesssim \|u\|_{L^\infty}^{r-2} \sum_{j\geq 1} G(j)^{br-1} \frac{2^{-j}}{\omega(2^{-j})} \left\| u_j^{\geq} \right\|_{L^2}^2 \\ &= \|u\|_{L^\infty}^{r-2} \|u\|_{\mathbb{B}^{\omega,a}}^2 ,\end{aligned}$$

which is the desired inequality raised to the r-th power. Note that the above proof works also for b) where the only small difference is that we use (78) to f_k instead of $f_k^{\leq}$ and then we chose the same f to get the same estimate as in a). We finally prove c). We start by applying Holder's inequality, first to the sum and then to the integral:

$$\begin{aligned}\left\| \left(\sum_{k\geq 1} G(k)^{2b} |u * \eta_k|^2 \right)^{\frac{1}{2}} \right\|_{L^r} &\leq \left\| \left(\sum_{k\geq 1} G(k)^{2a} |u * \eta_k|^2 \right)^{\frac{b}{2a}} \left(\sum_{k\geq 1} |u * \eta_k|^2 \right)^{\frac{a-b}{2a}} \right\|_{L^r} \\ &\leq \left\| \left(\sum_{k\geq 1} G(k)^{2a} |u * \eta_k|^2 \right)^{\frac{1}{2}} \right\|_{L^2}^{\frac{b}{a}} \left\| \left(\sum_{k\geq 1} |u * \eta_k|^2 \right)^{\frac{1}{2}} \right\|_{L^s}^{\frac{a-b}{a}} ,\end{aligned}$$

where $s = \dfrac{2(a-b)r}{2a-br}$ is finite under the assumption $b < \dfrac{2a}{r}$, which also implies $a > b$ since $r \geq 2$ by assumption. Using the Littlewood–Paley theorem (76), we obtain the following bound for the second factor:

$$\left\| \left(\sum_{k\geq 1} |u * \eta_k|^2 \right)^{\frac{1}{2}} \right\|_{L^s} \lesssim \|u\|_{L^s} \leq \|u\|_{L^\infty}.$$

We now use the L^2 almost-orthogonality of the Littlewood–Paley blocks, together with Parseval's identity, to estimate the remaining term

$$\begin{aligned}\left\| \left(\sum_{k\geq 1} G(k)^{2a} |u * \eta_k|^2 \right)^{\frac{1}{2}} \right\|_{L^2}^2 &= \int_{\mathbb{T}^d} \sum_{k\geq 1} G(k)^{2a} |u * \eta_k|^2 \\ &\leq \sum_{j\geq 1} \int_{\mathbb{T}^d} \sum_{k\geq 1} G(k)^{2a} |u_j * \eta_k|^2 \\ &= \sum_{j\geq 1} \sum_{k\geq 1} G(k)^{2a} \int_{\mathbb{T}^d} |u_j * \eta_k|^2 \\ &\lesssim \sum_{j\geq 1} \sum_{k\geq 1} G(k)^{2a} \sum_{\eta\in\mathbb{Z}^d} |\widehat{u}_j(\eta)|^2 |\widehat{\eta_k}(\eta)|^2 \\ &\leq \sum_{j\geq 1} \|u_j\|_{L^2}^2 \sum_{k\geq 1} G(k)^{2a} \sup_{\eta\in B_{2^j}(0)\setminus B_{2^{j-2}}(0)} |\widehat{\eta}_k(\eta)|^2.\end{aligned}$$

Let now $R > 0$ be such that the support of $\widehat{\eta_1}$ is contained in the ball $B_R(0)$. Note that, in order for the supremum above to be non-zero, it is necessary that $2^k R \geq 2^{j-2}$; otherwise, the support of $\widehat{\eta_k}$ does not intersect the annulus. We now use the fact that η_k has zero mean, i.e.,

$\widehat{\eta_k}(0) = 0$. By applying the multivariable Mean Value Theorem, we can estimate the supremum as follows:

$$\begin{aligned}\sup_{\eta\in B_{2^j}(0)\setminus B_{2^{j-2}}(0)} |\widehat{\eta}_k(\eta)|^2 &\leq 2^j \sup_{\xi\in B_{2^j}(0)} |\nabla\widehat{\eta}_k(\xi)| \\ &= 2^j \sup_{\xi\in B_{2^j}(0)} \left|\nabla\widehat{\eta}_1\left(\frac{\xi}{2^k}\right)\right| \cdot \frac{1}{2^k} \\ &= 2^{j-k} \sup_{\xi\in B_{2^j}(0)} \left|\nabla\widehat{\eta}_1\left(\frac{\xi}{2^k}\right)\right| \\ &\lesssim 2^{j-k}\|\nabla\widehat{\eta}_1\|_{L^\infty} \lesssim 2^{j-k}.\end{aligned}$$

Thanks to the previous computation, we are now interested in estimating the sum

$$\sum_{k\geq\lfloor j-2-\log_2 R\rfloor} G(k)^{2a}2^{j-k}.$$

In particular, it is not hard to see that this sum is bounded by a constant depending on R times $G(j)^{2a}$, thanks to the sublinear growth of G. With this observation, we are done since putting everything together we obtain

$$\left\|\left(\sum_{k\geq 1} G(k)^{2a}|u*\eta_k|^2\right)^{\frac{1}{2}}\right\|_{L^2} \lesssim \|u\|_{\mathbb{B}^{\omega,a}},$$

which implies the desired inequality, together with the fact that

$$\left\|\left(\sum_{k\geq 1} |u*\eta_k|^2\right)^{\frac{1}{2}}\right\|_{L^s} \lesssim \|u\|_{L^\infty}.$$

□

For the commutator estimate we will need a pointwise condition implied by the finiteness of (79).

Proposition 4.13 Let $f \in L^\infty(\mathbb{T}^d)$ be such that

$$\sum_{k=1}^{\infty} G(k)^{2a}\|f_k\|_{L^2}^2 < \infty.$$

Then there exists a function $g \in L^2(\mathbb{T}^d)$ such that for almost every $x, y \in \mathbb{T}^d$, we have

$$|f(x) - f(y)|G\left(-\log_2|x-y|\right)^a \leq g(x) + g(y).$$

Proof. It suffices to consider the case $|x - y| < 1$; otherwise, G is zero, and we can simply estimate g by setting $g = f$ and then taking the maximum with the function g constructed in the case $|x - y| < 1$. Fix k such that $2^{-k} \leq |x - y| < 2^{-k+1}$. Then, for almost every $x, y \in \mathbb{T}^d$, we have

$$|f(x) - f(y)| \leq |f_k^{\leq}(x) - f_k^{\leq}(y)| + |f_{k+1}^{\geq}(x) - f_{k+1}^{\geq}(y)|.$$

Moreover, by the same argument as in Lemma 4.9, we have that $G(x) \leq G(2x) \leq 2G(x)$. Therefore, it suffices to construct a function g such that

$$G(k)^a\left(|f_k^{\leq}(x) - f_k^{\leq}(y)| + |f_{k+1}^{\geq}(x) - f_{k+1}^{\geq}(y)|\right) \leq g(x) + g(y).$$

We first deal with the high-frequency component, using the triangle inequality:

$$G(k)^a|f_{k+1}^{\geq}(x) - f_{k+1}^{\geq}(y)| \leq G(k)^a\left(|f_{k+1}^{\geq}(x)| + |f_{k+1}^{\geq}(y)|\right).$$

Therefore, it suffices to show that the quantity

$$\sup_k G(k)^a|f_{k+1}^{\geq}(x)|$$

is bounded in $L^2(\mathbb{T}^d)$; but this is precisely what is established in (82).

We now turn to the remaining part. By (43), we have

$$|f_k^{\leq}(x) - f_k^{\leq}(y)| \lesssim 2^{-k}\left(M(|\nabla f_k^{\leq}|)(x) + M(|\nabla f_k^{\leq}|)(y)\right).$$

By the properties of the maximal function, it suffices to show that

$$\sup_k 2^{-k} G(k)^a |\nabla f_k^{\leq}(x)|$$

is bounded in $L^2(\mathbb{T}^d)$.

We observe that the square function

$$\left(\sum_{k=1}^{\infty}\left(2^{-k}G(k)^a|\nabla f_k^{\leq}(x)|\right)^2\right)^{1/2}$$

is an upper bound for the supremum, so it remains to show that this function belongs to $L^2(\mathbb{T}^d)$. In this final computation, we will use the fact that the L^2 norms of the gradients of the blocks are controlled by the L^2 norms of the blocks themselves up to a factor 2^k. We have the chain of equalities

$$\begin{aligned}\left\|\left(\sum_{k=1}^{\infty}\left(2^{-k}G(k)^a|\nabla f_k^{\leq}(x)|\right)^2\right)^{1/2}\right\|_{L^2}^2 &= \sum_{k=1}^{\infty}\left(2^{-k}G(k)^a\|\nabla f_k^{\leq}\|_{L^2}\right)^2\\ &\lesssim \sum_{k\geq 1}\sum_{1\leq j\leq k} 2^{-2k}G(k)^{2a}\|\nabla f_j\|_{L^2}^2\\ &\lesssim \sum_{j=1}^{\infty}\|f_j\|_{L^2}^2\sum_{n\geq j}G(n)^{2a}\frac{2^{-2n}}{2^{-2j}}\\ &= \sum_{j=1}^{\infty}\|f_j\|_{L^2}^2\sum_{n\geq j}G(n)^{2a}\frac{2^{2j}}{2^{2n}}.\end{aligned}$$

To end the proof, it suffices to show that

$$\sum_{n\geq j} G(n)^{2a}\cdot\frac{2^{2j}}{2^{2n}} \lesssim G(j)^{2a}.$$

We aim to provide a uniform finite upper bound for the sum on the left-hand side divided by $G(j)^{2a}$. Using the fact that $G(2x) \leq 2G(x)$, we split the sum into dyadic intervals:

$$\begin{aligned}\sum_{n\geq j}\frac{G(n)^{2a}}{G(j)^{2a}}\cdot\frac{2^{2j}}{2^{2n}} &= \sum_{j\leq n<2j}\frac{G(n)^{2a}}{G(j)^{2a}}\cdot\frac{2^{2j}}{2^{2n}} + \sum_{2j\leq n<4j}\frac{G(n)^{2a}}{G(j)^{2a}}\cdot\frac{2^{2j}}{2^{2n}} + \sum_{4j\leq n<8j}\cdots\\ &= \sum_{s=0}^{\infty}\sum_{n\in[2^s j, 2^{s+1}j)}\frac{G(n)^{2a}}{G(j)^{2a}}\cdot\frac{2^{2j}}{2^{2n}}.\end{aligned}$$

We now focus on a fixed dyadic block. In particular, for each $s \geq 0$, we estimate:

$$\sum_{2^s j\leq n<2^{s+1}j}\frac{G(n)^{2a}}{G(j)^{2a}}\cdot\frac{2^{2j}}{2^{2n}} \leq 2^{(s+1)2a}\cdot 2^{2j-2^{s+1}j} = 2^{(s+1)2a}\cdot 2^{(1-2^{s+1})j}.$$

Since $j \geq 1$, we have $2^{(1-2^{s+1})j} \leq 2^{1-2^{s+1}}$, and therefore:

$$\sum_{2^s j\leq n<2^{s+1}j}\frac{G(n)^{2a}}{G(j)^{2a}}\cdot\frac{2^{2j}}{2^{2n}} \leq 2^{(s+1)2a}\cdot 2^{1-2^{s+1}}.$$

Finally, summing over $s \geq 0$, we obtain:

$$\sum_{n\geq j}\frac{G(n)^{2a}}{G(j)^{2a}}\cdot\frac{2^{2j}}{2^{2n}} \lesssim \sum_{s=0}^{\infty} 2^{(s+1)2a}\cdot 2^{1-2^{s+1}} < \infty,$$

which proves the desired estimate. The sum is independent of j and clearly finite. □

Remark 4.14. Note that the previous proposition shows that, for a certain weight, a Littlewood–Paley integrability condition implies a pointwise condition like the one in (63). It was crucial in this proof that the growth of the weight $G(k)$ is sublinear, since the function G is close to a logarithm when ω is close to the identity.

The last part of this subsection will be related to discuss a similar question related to the condition (63). Note that when $\omega(x) = x$, the following three conditions are equivalent:

a) $f \in W^{1,p}$,

b) for almost every $x, y \in \mathbb{R}^d$, it holds that

$$|f(x) - f(y)| \leq |x - y|\big(g(x) + g(y)\big) \quad \text{for some } g \in L^p,$$

c)

$$\Big(\sum_k |f_k(x)\, 2^k|^2\Big)^{1/2} \in L^p.$$

One would like to understand how the previous conditions relate to their analogues when the modulus of continuity ω is of Osgood type. Note that, for our applications to PDE theory, it is not restrictive to assume that $\omega(x) \geq x$. Moreover, since ω is Osgood, it should not grow too fast — heuristically, it must be smaller than $x \log(1/x)^2$ as $x \to 0$, since this latter function is not an Osgood modulus. The question, therefore, is how the equivalence between the previous conditions behaves when ω is close to the identity.

In proving the previous equivalences, it was crucial several times that, at least in a weak sense, a gradient was available. In the present case, however, we can only define analogues of conditions **b)** and **c)**, or consider weighted quantities involving the gradient of the Littlewood–Paley blocks, in order to obtain quantities related to those previously considered in the ω-setting.

The natural replacement for the square function is clearly

$$\Big(\sum_k \Big|\frac{f_k(x)}{\omega(2^{-k})}\Big|^2\Big)^{1/2} \in L^p, \tag{85}$$

while the analogue of condition **b)** is clearly given by (63). It is not difficult to see that if $\omega(x)$ behaves like x^s for some $s \in (0, 1)$, then the square function condition in L^p implies the pointwise condition, but not vice versa. However, in our setting, the situation is more delicate. It is likely that the square function definition still implies the pointwise one, but when ω is close to $\omega(x) = x$, the converse might also hold. Under our assumptions on the vector field, we require both conditions to hold. In future work, we intend to investigate the relationship between these two conditions. We end this part with a proposition that bounds the supremum of a quantity involving the gradient of certain blocks. If one were able to prove that this supremum belongs to L^p under the square function definition instead of the pointwise one, by using the proof of Proposition 4.13 and some ideas from the appendix, it would follow that the square function definition implies the pointwise one.

Proposition 4.15 Let $p > 1$ and let $f \in L^p(\mathbb{R}^d)$, and suppose there exists $g \in L^p(\mathbb{T}^d)$ such that

$$|f(x) - f(y)| \leq \omega(|x - y|)\big(g(x) + g(y)\big) \qquad \text{for almost every } x, y \in \mathbb{T}^d. \tag{86}$$

Then,

$$\sup_k \frac{2^{-k}|\nabla f_k^{\leq}(x)|}{\omega(2^{-k})} \in L^p(\mathbb{T}^d).$$

Proof. The idea is in the spirit of Lemma 4.5. We prove a uniform bound in k on $\frac{2^{-k}|\nabla f_k^{\leq}(x)|}{\omega(2^{-k})}$ depending on g and f. By using the definition of convolution, how convolution interacts with integrals, and the fact that the gradient of a function has integral zero, we get

$$\frac{2^{-k}|\nabla f_k^{\leq}(x)|}{\omega(2^{-k})} = \int_{\mathbb{R}^d} \frac{2^{kd}}{\omega(2^{-k})}\big(f(y) - f(x)\big) * \nabla\psi(2^k|y - x|)\, dy.$$

We now divide the integral into dyadic annuli, starting from the ball of radius 2^{-k} centered at x. Inside the ball, by using (86), one can see that

$$\int_{B_{2^{-k}}(x)} \frac{2^{kd}}{\omega(2^{-k})}(f(y)-f(x)) * \nabla\psi(2^k|y-x|)\,dy \lesssim 2^{kd}\int_{B_{2^{-k}}(x)} (g(x)+g(y))\,dy \leq 2Mg(x).$$

Fix now $n \geq 1$ and consider the dyadic annulus $B_{2^{n-k}} \setminus B_{2^{n-k-1}}$, which we will denote by A_n for simplicity. By using (86) first, and then the fact that ω is concave, we get

$$\begin{aligned}
&\int_{A_n} \frac{2^{kd}}{\omega(2^{-k})}(f(y)-f(x)) * \nabla\psi(2^k|y-x|)\,dy \\
&\qquad \leq \int_{A_n} \frac{2^{kd}}{\omega(2^{-k})}(g(x)+g(y))\,\omega(2^{n-k})\,|\nabla\psi(2^k|y-x|)|\,dy \\
&\qquad \leq \int_{A_n} 2^{kd}\cdot 2^n\cdot (g(x)+g(y))\,|\nabla\psi(2^k|y-x|)|\,dy.
\end{aligned}$$

We now use the fact that ψ is a Schwartz function to obtain

$$\begin{aligned}
&\int_{A_n} 2^{kd}\cdot 2^n\cdot (g(x)+g(y))\,|\nabla\psi(2^k|y-x|)|\,dy \\
&\qquad \lesssim \int_{A_n} 2^{kd+n}(g(x)+g(y))\,\frac{1}{(2^k 2^{n-k})^{d+50}}\,dy \\
&\qquad = \int_{A_n} (g(x)+g(y))\,\frac{1}{2^{n\cdot 50}}\,dy \\
&\qquad \lesssim \frac{1}{2^n} \fint_{B_{2^{k-n}}(x)} (g(x)+g(y))\,dy \\
&\qquad \lesssim \frac{Mg(x)}{2^n}.
\end{aligned}$$

This concludes the proof by summing over all scales $n \geq 1$. □

3.3. Main Regularity Estimate. In this section, we prove the main regularity estimate for solutions of the transport equation. From now on, the modulus of continuity ω is assumed to satisfy the following technical condition:

$$\sum_{n=1}^{\infty} \sup_{k\geq 1} \left| \frac{\omega\left(2^{-(k+n)}\right)}{\omega\left(2^{-k}\right)} \right| < \infty. \tag{87}$$

Remark 4.16. This condition is quite natural and is satisfied by all standard examples of moduli of continuity, such as $x\log(\frac{1}{x})^{\beta}$.

As said in the previous subsection, we will assume both the square function definition and the pointwise one. More precisely, we require that for almost every $t \in [0,T]$ we have

$$s_t = \sqrt{\sum_{k=1}^{\infty} \left| \frac{b_k(t,\cdot)}{\omega(2^{-k})} \right|^2} \in L^p,$$

and

$$\int_0^T \|s_t\|_{L^p}\,dt < \infty,$$

and moreover $b \in L^1([0,T]; W^{\omega,p}(\mathbb{T}^d))$.

Furthermore, we define

$$\tilde{g}_t = \max\{s_t, g_t\}$$

and we note that

$$\int_0^T \|\tilde{g}_t\|_{L^p}\,dt \lesssim \int_0^T (\|g_t\|_{L^p} + \|s_t\|_{L^p})\,dt.$$

We will prove the following theorem.

Theorem 4.17 Let $p \in (1, \infty)$ be given and let be a divergence-free vector field satisfying the pointwise and square function conditions. Let $u \in L^\infty([0,T] \times \mathbb{T}^d)$ be a mean-free solution to the transport equation. If $u_0 \in B^{\omega,a}(\mathbb{T}^d)$ for some $\frac{1}{2} \le a < \frac{p}{2}$, then

$$u \in L^\infty_{\mathrm{loc}}([0,T]; B^{\omega,a}(\mathbb{T}^d)),$$

and there exists a constant C, depending on d, p, and a, such that

$$\|u(t)\|_{B^{\omega,a}} \le C \left(\int_0^t \|\tilde{g}_s\|_{L^p} \, ds \right)^a \|u\|_{L^\infty} + C\|u_0\|_{B^{\omega,a}},$$

for any $t \in [0,T]$.

Remark 4.18. Note that requiring the solution to be mean-free is natural, since the mean is preserved in time when the flow is divergence-free. Indeed, testing the equation with a function of the form $\varphi(t,x) = \psi(t)$, where $\psi \in C_c^\infty([0,T))$, yields that the distributional time derivative of the function

$$t \mapsto \int_{\mathbb{T}^d} u(t,x) \, dx$$

is zero. Hence, the mean value of u is constant in time.

Thanks to the previous theorem, we will be able to prove in the final subsection the following result on the renormalization property of solutions to the transport equation.

Theorem 4.19 Let b be a bounded and divergence-free vector field that satisfies the integrability hypotheses associated with the Osgood modulus of continuity ω, which satisfies (87). Then, bounded solutions of the transport equation with initial datum in $L^\infty \cap B^{\omega,a}$ are renormalized if $p \ge 2$ and

$$\frac{\omega(\delta)}{\delta \left(\int_\delta^1 \frac{ds}{\omega(s)} \right)^a} \to 0 \tag{88}$$

for some $a < \frac{p}{2}$ as $\delta \to 0$.

Remark 4.20. Note that, for example, $\omega(x) = x\sqrt{\log \frac{1}{x}}$ satisfies (93) whenever $p > 2$.

The rest of this subsection is devoted to the proof of the theorem. We aim to derive a transport equation for each component u_k, where the right-hand side will contain a commutator term to account for the fact that only the full function u satisfies the transport equation. We will then use this equation to estimate the time derivative of the L^2 norm at a given time by testing the distributional equation against a suitable Littlewood–Paley block. From now on, the time variable t will be left implicit in all expressions, as it will not play a significant role in the argument.

As in the computations from Chapter 2, namely (26) and (27), we observe that (since the Littlewood–Paley blocks are defined via convolution) the function $u_k^{\geq}$ satisfies the equation

$$\partial_t u_k^{\geq} + b \cdot \nabla u_k^{\geq} = r_k,$$

where r_k is a commutator term. By performing the same computation as before, and using the fact that u is a solution, we find, as in the proof of (28), that

$$r_k = \int_{\mathbb{R}^d} u(x-y)(b(x) - b(x-y)) \cdot \nabla\psi_{k-1}(y) \, dy.$$

We now test the equation with u_j for $j \ge k$. Since the commutator is bounded in the L^∞ norm, we can sum the equations over $j \ge k$, and using the fact that $f_k^{\geq} = \sum_{j \ge k} f_j$ in L^p, we deduce

$$\frac{1}{2} \frac{d}{dt} \int_{\mathbb{T}^d} |u_k^{\geq}|^2 = \int_{\mathbb{T}^d} u_k^{\geq} \, r_k.$$

This last expression can be rewritten as

$$\frac{d}{dt} \int_{\mathbb{T}^d} |u_k^{\geq}|^2 \, dx = \int_{\mathbb{T}^d} \int_{\mathbb{R}^d} u_k^{\geq}(x) \, u(x-y) \, (b(x) - b(x-y)) \cdot \nabla\psi_{k-1}(y) \, dy \, dx$$

We now aim to rewrite the right-hand side of the inequality more clearly by identifying which Littlewood–Paley blocks interact. We start from the identity

$$\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(x)\, u(x-y)\, b(x)\cdot\nabla\psi_{k-1}(y)\, dy\, dx = \int_{\mathbb{T}^d} u_k^{\geq}(x)\, b(x)\cdot\big(u*\nabla\psi_{k-1}\big)(x)\, dx.$$

We now decompose u and b using their Littlewood–Paley decompositions. In particular, we want to understand, for fixed $m, n \geq 1$, the contribution of the term

$$\int_{\mathbb{T}^d} u_k^{\geq}(x)\, b_m(x)\cdot\big(u_n*\nabla\psi_{k-1}\big)(x)\, dx.$$

By using Parseval's identity and the way convolution and products behave under the Fourier transform, we obtain

$$\int_{\mathbb{T}^d} u_k^{\geq}(x)\, b_m(x)\cdot(u_n*\nabla\psi_{k-1})(x)\, dx = \sum_{\eta'\in\mathbb{Z}^d} \widehat{u_k^{\geq} b_m}(-\eta')\cdot\widehat{u_n}(\eta')\,\widehat{\nabla\psi_{k-1}}(\eta').$$

Using the identity

$$\widehat{u_k^{\geq} b_m}(-\eta') = \sum_{\eta\in\mathbb{Z}^d} \widehat{u_k^{\geq}}(-\eta)\,\widehat{b_m}(\eta-\eta'),$$

we rewrite the sum as

$$\sum_{\eta,\eta'\in\mathbb{Z}^d} \widehat{u_k^{\geq}}(-\eta)\,\widehat{b_m}(\eta-\eta')\cdot\widehat{u_n}(\eta')\,\widehat{\nabla\psi_{k-1}}(\eta').$$

We now want to understand, for a fixed k, for which indices n, m the above sum is zero. We observe that, in order for the sum to be nonzero, we must have $k-2 < n \leq k$; otherwise, $\widehat{\nabla\psi_{k-1}}(\eta')$ and $\widehat{u_n}(\eta')$ have disjoint supports.

Similarly, for a nonzero contribution to the sum, we must have $|\eta|2^{k-2}$ and $2^{n-2} < |\eta'| \leq 2^n$. Moreover, since we need $2^{m-2} \leq |\eta-\eta'| \leq 2^m$, we deduce that

$$2^{k-2} \leq |\eta| \leq |\eta'| + |\eta-\eta'| \leq 2^n + 2^m.$$

We now turn to the second part of the integral, i.e.,

$$-\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(x) u(x-y) b(x-y)\cdot\nabla\psi_{k-1}(y)\, dy\, dx.$$

We proceed as above, i.e., we expand in Littlewood-Paley blocks. Thus, we are interested, as before, in determining for which indices n, m the integral is nonzero. By a similar computation, passing to the Fourier transform, we have

$$\begin{aligned}
&\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(x) u_n(x-y) b_m(x-y)\cdot\nabla\psi_{k-1}(y)\, dy\, dx \\
&= \int_{\mathbb{T}^d} u_k^{\geq}(x)\big(\nabla\psi_{k-1}*(b_m u_n)\big)(x)\, dx \\
&= \sum_{\eta\in\mathbb{Z}^d} \widehat{u_k^{\geq}}(-\eta)\,\widehat{\nabla\psi_{k-1}*(b_m u_n)}(\eta) \\
&= \sum_{\eta,\eta'\in\mathbb{Z}^d} \widehat{u_k^{\geq}}(-\eta)\,\widehat{\nabla\psi_{k-1}}(\eta)\cdot\widehat{b_m}(\eta')\widehat{u_n}(\eta-\eta').
\end{aligned}$$

We now argue as before and observe that we must have

$$2^{k-2} \leq |\eta| \leq 2^{k-1}, \quad 2^{m-2} \leq |\eta'| \leq 2^m, \quad \text{and} \quad 2^{n-2} \leq |\eta-\eta'| \leq |\eta| + |\eta'| \leq 2^{k-1} + 2^m.$$

By inspecting the conditions, we notice that both contributions vanish if

$$n > k+4 \quad \text{and} \quad m < k+3.$$

Thus, recalling the definition of the norm and the equation given before, we have that

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|^2_{B^{\omega,a}} &= \sum_{k\geq 1} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(x)\,u(x-y)\,b^{\geq}_{k+3}(x)\cdot\nabla\psi_{k-1}(y)\,\mathrm{d}y\,\mathrm{d}x\\
&- \sum_{k\geq 1} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(x)\,u(x-y)\,b^{\geq}_{k+3}(x-y)\cdot\nabla\psi_{k-1}(y)\,\mathrm{d}y\,\mathrm{d}x\\
&+ \sum_{k\geq 1} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(x)\,u^{\leq}_{k+4}(x-y)\left(b^{\leq}_{k+2}(x)-b^{\leq}_{k+2}(x-y)\right)\cdot\nabla\psi_{k-1}(y)\,\mathrm{d}y\,\mathrm{d}x\\
&=: \mathrm{I}-\mathrm{II}+\mathrm{III}.\end{aligned}$$

From now on, we will often write α instead of $2a-1$. We are going to show that I, II and III can all be estimated by

$$|\mathrm{I}|+|\mathrm{II}|+|\mathrm{III}| \lesssim \|u\|_{L^\infty}^{\frac{1}{a}}\,\|u\|_{B^{\omega,a}}^{\frac{2a-1}{a}}\,\|\tilde g\|_{L^p}. \tag{89}$$

This last estimate yields the desired inequality in the thesis, since it implies that

$$a\frac{d}{dt}\|u\|^{1/a}_{B^{\omega,a}} \lesssim \|\tilde g\|_{L^p}\|u\|^{1/a}_{L^\infty},$$

which, once integrated in time, gives exactly the desired bound. We start by estimating I. We will divide $\alpha=\alpha_1+\alpha_2$ for α_1,α_2 to be chosen later. Similarly, let q,s satisfy

$$\frac{1}{p}+\frac{1}{q}+\frac{1}{s}=1.$$

Splitting $b^{\geq}_{k+3}$ into phase blocks, we may write

$$\mathrm{I}=\sum_{n\geq 3}^{\infty}\sum_{k\geq 1}^{\infty}\frac{\omega(2^{-(k+n)})}{\omega(2^{-k})}\int_{\mathbb{T}^d} G(k)^{\alpha}u_k^{\geq}\left(\frac{1}{\omega(2^{-(k+n)})}b_{k+n}\right)\cdot\left(u*(2^{-k}\nabla\psi_{k-1})\right)dx. \tag{90}$$

Using the convergence property of the series in ω, i.e.(87), it suffices to estimate the sum over k uniformly in n. Invoking the Cauchy–Schwarz inequality for the sums and subsequently the Hölder inequality for the integrals, we find that

$$\begin{aligned}|\mathrm{I}| &\leq \int_{\mathbb{T}^d}\left(\sup_{k\geq 1}G(k)^{\alpha_1}|u_k^{\geq}|\right)\left(\sum_{k\geq 1}\Big|\frac{1}{\omega(2^{-(k+n)})}b_{k+n}\Big|^2\right)^{\frac12}\left(\sum_{k\geq 1}G(k)^{2\alpha_2}\big|u*(2^{-k}\nabla\psi_{k-1})\big|^2\right)^{\frac12}dx\\
&\leq \left\|\sup_k G(k)^{\alpha_1}|u_k^{\geq}|\right\|_{L^q}\left\|\left(\sum_{k\geq 1}\Big|\frac{1}{\omega(2^{-(k+n)})}b_{k+n}\Big|^2\right)^{\frac12}\right\|_{L^p}\left\|\left(\sum_{k\geq 1}G(k)^{2\alpha_2}\big|u*(2^{-k}\nabla\psi_{k-1})\big|^2\right)^{\frac12}\right\|_{L^s}.\end{aligned}$$

To conclude the estimate in this case, we use (82), (84) and estimate the second term with the whole square function related to the vector field. More precisely, we apply (82) with $b=\alpha_1=\frac{\alpha}{2}$ and $r=q=\frac{2p}{p-1}$ to the first term, and (84) with $r=s=\frac{2p}{p-1}$ and $b=\alpha_2=\frac{\alpha}{2}$ to the third term.

These choices are admissible provided that

$$\frac{\alpha p}{p-1}<2a,$$

which is equivalent to the condition $a<\frac{p}{2}$. In particular, we obtain an estimate of the form (89).

We now turn to proving a similar bound for the term II. We start with the expression

$$\sum_{k\geq 1} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(x)\,u(x-y)\,b^{\geq}_{k+3}(x-y)\cdot\nabla\psi_{k-1}(y)\,\mathrm{d}y\,\mathrm{d}x.$$

We now perform the change of variables $z=x-y$that gives

$$\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u_k^{\geq}(z+y)\,u(z)\,b^{\geq}_{k+3}(z)\cdot\nabla\psi_{k-1}(y)\,\mathrm{d}y\,\mathrm{d}z.$$

This can be rewritten as

$$\int_{\mathbb{T}^d} u(z)\, b^{\geq}_{k+3}(z) \cdot \left(u^{\geq}_k * \nabla\psi_{k-1}\right)(z)\,\mathrm{d}z$$

since $\nabla\psi_{k-1}$ is odd(Littlewood-Paley use radial kernels).

Plugging this back into the sum gives the identity:

$$\begin{aligned}\sum_{k\geq 1} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} u^{\geq}_k(x)\,u(x-y)\,b^{\geq}_{k+3}(x-y)\cdot\nabla\psi_{k-1}(y)\,\mathrm{d}y\,\mathrm{d}x \\ = -\sum_{k\geq 1} G(k)^{2a-1}\frac{2^{-k}}{\omega(2^{-k})}\int_{\mathbb{T}^d} u\, b^{\geq}_{k+3}\cdot\left(u^{\geq}_k * \nabla\psi_{k-1}\right)\,\mathrm{d}x.\end{aligned}$$

Note now that, by construction, we have

$$\left(u^{\geq}_k * \nabla\psi_{k-1}\right) = \left(u_k * \nabla\psi_{k-1}\right),$$

since the support of $\nabla\psi_{k-1}$ is concentrated at frequencies $\leq 2^{k-1}$, which do not interact with the high-frequency part of $u^{\geq}_k$. Moreover, by a similar argument (namely, applying Parseval's identity and observing that the Fourier supports are disjoint), one can see that the low-frequency part $u^{\leq}_k$ of u does not contribute. This is because the Fourier transform of $b^{\geq}_{k+3}$ vanishes inside the ball of radius 2^{k+1} centered at the origin, where the Fourier transform of $u^{\leq}_k$ is supported. As a result,

$$\int_{\mathbb{T}^d} u\, b^{\geq}_{k+3}\cdot(u^{\geq}_k * \nabla\psi_{k-1})\,dx = \int_{\mathbb{T}^d} u^{\geq}_{k+1}\, b^{\geq}_{k+3}\cdot(u^{\geq}_k * \nabla\psi_{k-1})\,dx,$$

since we can replace u with $u^{\geq}_{k+1}$ in the integral without changing its value. Thus, we obtain

$$\mathrm{II} = -\sum_{k\geq 1} G(k)^{\alpha}\frac{2^{-k}}{\omega(2^{-k})}\int_{\mathbb{T}^d} u^{\geq}_{k+1}\, b^{\geq}_{k+3}\cdot(u^{\geq}_k * \nabla\psi_{k-1})\,dx.$$

This last expression is essentially the same as I and can be handled in exactly the same way. Finally we prove the estimate for III. By using the fundamental theorem of calculus we rewrite this term as

$$\begin{aligned}&\int_{\mathbb{R}^d} u^{\leq}_{k+4}(x-y)(b^{\leq}_{k+2}(x) - b^{\leq}_{k+2}(x-y))\nabla\psi_{k-1}(y)\,dy \\ &= \int_0^1\int_{\mathbb{R}^d} u^{\leq}_{k+4}(x-y)\,\nabla b^{\leq}_{k+2}(x-sy) : \nabla\psi_{k-1}(y)\otimes y\,dy\,ds \\ &= \int_0^1\int_{\mathbb{R}^d} u^{\leq}_{k+4}(x-y)\left(\nabla b^{\leq}_{k+2}(x-sy) - \nabla b^{\leq}_{k+2}(x)\right) : \nabla\psi_{k-1}(y)\otimes y\,dy\,ds \\ &\qquad + \nabla b^{\leq}_{k+2}(x) : \int_{\mathbb{R}^d} u^{\leq}_{k+4}(x-y)\,\nabla\psi_{k-1}(y)\otimes y\,dy \\ &=: h_k(x) + p_k(x)\end{aligned}$$

where $\otimes$ denotes the *tensor product*, defined by

$$(a\otimes b)_{ij} = a_i b_j \quad \text{for } a,b\in\mathbb{R}^d,$$

and : denotes the *Frobenius inner product* (also called double contraction), given by

$$A : B = \sum_{i,j=1}^d A_{ij}B_{ij} = \mathrm{Tr}(A^T B) \quad \text{for } A,B\in\mathbb{R}^{d\times d}.$$

We will deal with h_k and p_k separately. We start from h_k, where after noticing that

$$\begin{aligned}&\int_{\mathbb{R}^d}\left(\nabla b^{\leq}_{k+2}(x-sy) - \nabla b^{\leq}_{k+2}(x)\right) : \nabla\psi_{k-1}(y)\otimes y\,dy\,ds \\ &= \int_{\mathbb{R}^d}\psi_{k-1}\,\mathrm{div}_y(y\cdot\nabla b^{\leq}_{k+2}(x-sy)) - \langle \mathrm{div}_y\, b^{\leq}_{k+2}(x-sy), \nabla\psi_{k-1}(y)\rangle\,dy\,ds = 0,\end{aligned}$$

we insert an extra factor $u^{\leq}_{k+4}(x)$ in h_k and write

$$h_k(x) = \int_0^1 \int_{\mathbb{R}^d} \Big(u^{\leq}_{k+4}(x-y) - u^{\leq}_{k+4}(x)\Big) \times \Big(\nabla b^{\leq}_{k+2}(x-sy) - \nabla b^{\leq}_{k+2}(x)\Big) : \nabla\psi_{k-1}(y) \otimes y \, dy \, ds.$$

We can further manipulate this by expressing the differences in terms of mean values:

$$h_k(x) = \int_0^1 \int_0^1 \int_0^1 \int_{\mathbb{R}^d} y \cdot \nabla u^{\leq}_{k+4}(x-ry) \nabla^2 b^{\leq}_{k+2}(x-sty) : \nabla\psi_{k-1}(y) \otimes y \otimes y \, dy \, dr \, ds \, dt.$$

Splitting u and b into phase blocks, and using the convention that $u_\ell = 0$ and $b_m = 0$ for $\ell, m \leq 0$, the latter is furthermore controlled by

$$\begin{aligned} |h_k(x)| &\leq \sum_{j\geq -4} \sum_{n \geq -2} \int_0^1 \int_0^1 \int_0^1 \int_{\mathbb{R}^d} |\nabla u_{k-j}(x-ry)| \\ &\qquad \times |\nabla^2 b_{k-n}(x-sty)| \, |\nabla\psi_{k-1}(y)| \, |y|^3 \, dy \, dr \, ds \, dt \\ &= \sum_{j \geq -4} 2^{-j} \sum_{n \geq -2} \int_0^1 \int_0^1 \int_0^1 \int_{\mathbb{R}^d} 2^{j-k} |\nabla u_{k-j}(x-ry)| \\ &\qquad \times 2^{-k} \frac{\omega(2^{n-k}) 2^{2(k-n)}}{\omega(2^{n-k}) 2^{2(k-n)}} |\nabla^2 b_{k-n}(x-sty)| \, \rho_k(y) \, dy \, dr \, ds \, dt, \end{aligned}$$

where we have introduced

$$\rho_k(y) := 4^k |y|^3 |\nabla\psi_{k-1}(y)|$$

for notational convenience.

We want to argue in a similar way as in I. In particular, fixing n, we want to estimate the quantity

$$\frac{2^{-k}}{\omega(2^{-k})} \cdot 2^{-k} \cdot \omega(2^{n-k}) \cdot 2^{2(n-k)}.$$

Using the fact that for $y \leq 1$ we have, by the fact that ω is concave, that $\omega(x)y \leq \omega(xy)$, we conclude that the whole expression is bounded by 2^{-n}. Multiplying by $G(k)^{\alpha_2} \frac{2^{-k}}{\omega(2^{-k})}$, summing over k, and using Holder inequality, we obtain

$$\begin{aligned} \sum_{k \geq 1} G(k)^{\alpha_2} \frac{2^{-k}}{\omega(2^{-k})} |h_k(x)| &\leq \sum_{j \geq -4} 2^{-j} \sum_{n \geq -2} 2^{-n} \int_0^1 \int_0^1 \int_0^1 \int_{\mathbb{R}^d} \left(\sum_{k \geq 1} G(k)^{2\alpha_2} 2^{2(j-k)} |\nabla u_{k-j}(x-ry)|^2 \right)^{\frac{1}{2}} \\ &\qquad \times \left(\sum_{k \geq 1} \frac{2^{4(n-k)}}{\omega(2^{n-k})^2} |\nabla^2 b_{k-n}(x-sty)|^2 \right)^{\frac{1}{2}} \rho_k(y) \, dy \, dr \, ds \, dt. \end{aligned}$$

Therefore, integrating against $G(k)^{\alpha_1} u^{\geq}_k$ and applying Hölder's inequality inside the integrals, we deduce that

$$\begin{aligned} \sum_{k \geq 1} G(k)^{\alpha} \frac{2^{-k}}{\omega(2^{-k})} \int_{\mathbb{T}^d} |u^{\geq}_k(x)| |h_k(x)| \, dx &\leq \sum_{j \geq -4} 2^{-j} \sum_{n \geq -2} 2^{-n} \int_0^1 \int_0^1 \int_0^1 \int_{\mathbb{R}^d} \left\| \sup_{k \geq 1} G(k)^{\alpha_1} |u^{\geq}_k| \right\|_{L^q} \\ &\qquad \times \left\| \left(\sum_{k \geq 1} G(k)^{2\alpha_2} 2^{2(j-k)} |\nabla u_{k-j}(\cdot - ry)|^2 \right)^{\frac{1}{2}} \right\|_{L^s} \\ &\qquad \cdot \left\| \left(\sum_{k \geq 1} \frac{2^{4(n-k)}}{\omega(2^{n-k})^2} |\nabla^2 b_{k-n}(\cdot - sty)|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \rho_k(y) \, dy \, dr \, ds \, dt. \end{aligned}$$

We now make use of the periodicity of the problem and the convergence of geometric series to deduce

$$\sum_{k\geq 1} G(k)^\alpha \int_{\mathbb{T}^d} |u_k^{\geq}(x)||h_k(x)|\,dx \leq \left\|\sup_{k\geq 1} G(k)^{\alpha_1}|u_k^{\geq}|\right\|_{L^q} \left\|\left(\sum_{\ell\geq 1} G(\ell)^{2\alpha_2}2^{-2\ell}|\nabla u_\ell|^2\right)^{\frac{1}{2}}\right\|_{L^s}$$
$$\times \left\|\left(\sum_{m\geq 1}\frac{2^{-4m}}{\omega(2^{-m})^2}|\nabla^2 b_m|^2\right)^{\frac{1}{2}}\right\|_{L^p} \|\rho_k\|_{L^1}.$$

In order to obtain an estimate as in (89), we conclude with a few observations. The first two terms are handled using the usual interpolation inequalities from the lemma, after noting that $2^{-\ell}\nabla u_\ell = u * \eta_\ell$, where $\eta_\ell = 2^{-\ell}\nabla\varphi_\ell$ satisfies the assumptions of (84). The third term follows directly from (99), while the fourth term is bounded by observing that $\|\rho_k\|_{L^1} \sim \|\rho_1\|_{L^1} \sim 1$, due to scaling.

We finally deal with the term p_k. We decompose

$$\nabla\psi_{k-1}\otimes y = \Phi_k - I\psi_{k-1}, \quad \text{where } \Phi_k := \nabla(y\psi_{k-1}), \tag{91}$$

and observe that $\nabla b_{k+2}^{\leq} : I = \operatorname{div} b_{k+2}^{\leq} = 0$ by the incompressibility assumption on b. Moreover, Φ_k is an even function. It follows that we may rewrite

$$p_k(x) = \nabla b_{k+2}^{\leq}(x) : (u_{k+4}^{\leq} * \Phi_k)(x).$$

Thanks to the fact that the Fourier transform of Φ_k is supported in the annulus $B_{2^{k-1}}(0) \setminus B_{2^{k-2}}(0)$, we have

$$\theta * \Phi_k = (\theta_k + \theta_{k-1}) * \Phi_k.$$

Inserting this identity into the full integral, we obtain

$$\sum_{k\geq 1} G(k)^\alpha \int_{\mathbb{T}^d} \frac{2^{-k}}{\omega(2^{-k})}\theta_k^{\geq} g_k^2\,dx = \sum_{k\geq 1} G(k)^\alpha \int_{\mathbb{T}^d} \frac{2^{-k}}{\omega(2^{-k})} u_k^{\geq}\,\nabla b_{k+2}^{\leq} : \left(u * \eta_{k-1}^1 * \Phi_k\right)\,dx, \tag{92}$$

where we have introduced the mollifiers $\eta_k^j := \phi_k + \cdots + \phi_{k+j}$. We may now apply Parseval's identity in a similar manner as before, and observe that only the frequencies of $u_k^{\geq}$ smaller than 2^{k+3} contribute to the integral. Hence, we may replace $u_k^{\geq}$ by $u * \eta_k^4$. We thus arrive at

$$\sum_{k\geq 1} G(k)^\alpha \int_{\mathbb{T}^d} \frac{2^{-k}}{\omega(2^{-k})} u_k^{\geq} p_k\,dx = \sum_{k\geq 1} G(k)^\alpha \int_{\mathbb{T}^d} \frac{2^{-k}}{\omega(2^{-k})} \left(u * \eta_k^4\right)\,\nabla b_{k+2}^{\leq} : \left(u * \eta_{k-1}^1 * \Phi_k\right)\,dx.$$

We want to bound the supremum over k of the quantity

$$\frac{2^{-k}}{\omega(2^{-k})}|\nabla b_{k+2}^{\leq}(x)|$$

by an L^p-function $s(x)$ which we can control and this is done by proposition 4.15.

We conclude as follows. Applying the Cauchy–Schwarz inequality to the sum, followed by Hölder's inequality in the integral, we estimate the first term by

$$\int_{\mathbb{T}^d} |s(x)| \left(\sum_{k\geq 1} G(k)^\alpha |u * \eta_k^4|^2\right)^{\frac{1}{2}} \left(\sum_{k\geq 1} G(k)^\alpha |u * \eta_{k-1}^1 * \Phi_k|^2\right)^{\frac{1}{2}} dx.$$

This is bounded by

$$\|s\|_{L^p} \left\|\left(\sum_{k\geq 1} G(k)^\alpha |u * \eta_k^4|^2\right)^{\frac{1}{2}}\right\|_{L^q} \left\|\left(\sum_{k\geq 1} G(k)^\alpha |u * \eta_{k-1}^1 * \Phi_k|^2\right)^{\frac{1}{2}}\right\|_{L^s}$$

which is bounded by the bound in (89) by using the usual interpolation inequality (84) and the fact that $\|s\|_{L^p} \lesssim \|\tilde{g}\|_{L^p}$.

3.4. Commutator estimate. In this last subsection, we use the previously established regularity to prove the renormalization property for solutions of the transport equation. We assume the hypotheses of the regularity theorem hold. In particular, we have that there exist $g_1 \in L^2$, $g_2 \in L^p$

$$|u(x) - u(y)| \lesssim \frac{g_1(x) + g_1(y)}{G(-\log_2 |x-y|)^a}$$

for almost every $x, y \in \mathbb{T}^d$, and

$$|b(x) - b(y)| \lesssim \omega(|x-y|)\big(g_2(x) + g_2(y)\big)$$

for almost every $x, y \in \mathbb{T}^d$. We remember, as seen in Chapter 2, that in order to have the renormalization property it suffices to show that the commutator goes to zero in L^1_{loc}. The commutator is given by (since ρ has support in the unit ball)

$$r_\delta(u,b)(x) = \int_{|h|\le 1} (u(x) - u(x-\delta h)) \frac{b(x) - b(x-\delta h)}{\delta} \cdot \nabla\rho(h)\, dh,$$

where the extra term compared to (28) arises from the fact that b is divergence-free. Since we want to take the limit $\delta \to 0$, we consider δ very small.

Using the two pointwise bounds and bounding the mollifier, we obtain

$$|r_\delta(u,b)(x)| \lesssim \int_{|h|\le 1} \frac{\omega(\delta|h|)}{\delta\big(\int_{\delta|h|}^1 \frac{ds}{\omega(s)}\big)^a} \big(p(x) + p(x-\delta h)\big)\big(g(x) + g(x-\delta h)\big) dh.$$

Moreover, since $|h| \le 1$, we have

$$\frac{\omega(\delta|h|)}{\delta\big(\int_{\delta|h|}^1 \frac{ds}{\omega(s)}\big)^a} \le \frac{\omega(\delta)}{\delta\big(\int_{\delta}^1 \frac{ds}{\omega(s)}\big)^a}.$$

We now consider

$$\int_{|h|\le 1} \int_{\mathbb{T}^d} (p(x) + p(x-\delta h))(g(x) + g(x-\delta h))\, dx\, dh.$$

By applying the Cauchy–Schwarz inequality with respect to the integration over the torus and using the fact that the torus has finite measure, it is easy to see that the integral is bounded uniformly in δ, provided that

$$\frac{1}{2} + \frac{1}{p} \le 1.$$

In particular, if $p \ge 2$ and

$$\frac{\omega(\delta)}{\delta\left(\int_\delta^1 \frac{ds}{\omega(s)}\right)^a} \to 0 \tag{93}$$

for some $a < \frac{p}{2}$, as $\delta \to 0$, then every bounded solution of the equation (satisfying a mild regularity condition) is renormalized.

Appendix

A.1. Real analysis tools

This section gathers several results used mainly in the proof of the Ambrosio Superposition Principle (Theorem 2.17).

Let $\mathcal{M}_+(X)$ denote the set of all nonnegative locally finite Borel measures on X.

A family $\mathcal{F} \subset \mathcal{M}_+(X)$ is said to be *bounded* if there exists a constant $C > 0$ such that

$$\|\mu\|_{\mathcal{M}(X)} \leq C \quad \text{for every } \mu \in \mathcal{F}.$$

We say that a bounded sequence $\{\mu_h\} \subset \mathcal{M}_+(X)$ *converges narrowly* to $\mu \in \mathcal{M}_+(X)$ as $h \to \infty$ if

$$\lim_{h\to\infty} \int_X f(x)\, d\mu_h(x) = \int_X f(x)\, d\mu(x)$$

for every $f \in C_b(X)$, the space of continuous and bounded real-valued functions on X.

A bounded family $\mathcal{F} \subset \mathcal{M}_+(X)$ is said to be *tight* if for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset X$ such that

$$\mu(X \setminus K_\varepsilon) \leq \varepsilon \quad \text{for every } \mu \in \mathcal{F}.$$

The following theorem relates tightness to compactness. For a more general discussion, see [**Bil99**], which provides a comprehensive treatment of the convergence of probability measures. The arguments can be adapted to the case of finite measures.

Theorem A.21 (Prokhorov) Assume that X is a complete metric space. Then a bounded family $\mathcal{F} \subset \mathcal{M}_+(X)$ is relatively compact with respect to narrow convergence if and only if it is tight.

Coercive functionals are closely related to the Prokhorov theorem, since they provide a useful criterion to ensure tightness.

Definition A.22. Assume that X is a complete metric space. A (lower semi-continuous) functional

$$\Psi : X \to [0, \infty]$$

is said to be coercive if all its sub-level sets $\{x \in X : \Psi(x) \leq c\}$, for $c \in \mathbb{R}$, are compact.

In this setting, the following criterion holds.

Theorem A.23 (Tightness via coercive integrands) Let X be a complete metric space and let $\mathcal{F} \subset \mathcal{P}(X)$ be a family of probability measures. Then the family $\mathcal{F}$ is tight if and only if there exists a (lower semi-continuous) coercive function $\Psi : X \to [0, \infty]$ such that

$$\sup_{\mu\in\mathcal{F}} \int_X \Psi(x)\, d\mu(x) < \infty.$$

We conclude with a basic fact about tightness. A bounded family $F \subset \mathcal{M}_+(X \times Y)$ is tight if and only if the families of its marginals

$$\pi_X(F) \subset \mathcal{M}_+(X) \quad \text{and} \quad \pi_Y(F) \subset \mathcal{M}_+(Y)$$

are tight.

We conclude this section by stating a special case of the Disintegration Theorem, which suffices for our purposes. For a proof and a more general treatment, see Section 2.5 of [**AFP00**].

Theorem A.24 Let $\mu \in \mathcal{M}_+(X \times Y)$, and define $\nu = (\pi_X)_\# \mu$, assuming that $\nu \in \mathcal{M}_+(X)$. Then there exists a Borel family $\{\mu_x\}_{x \in X} \subset \mathcal{M}_+(Y)$, uniquely determined ν-almost everywhere, such that

$$\mu = \int_X \mu_x \, d\nu(x).$$

A.2. Standard Harmonic analysis tools

We write $A \lesssim B$ to mean that $A \leq CB$ for some constant C, often depending only on the dimension of the space. If the constant depends on other parameters, this will either be clear from the context or explicitly stated. Moreover, we sometimes use $|X|$ to denote $\mathscr{L}^d(X)$.

In this section, we introduce the Hardy–Littlewood maximal function and we present some properties.

Definition A.25. Given a measurable function f, the Hardy-Littlewood maximal function is defined as

$$M(f)(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)| \, dy, \tag{94}$$

where the supremum is taken over all balls $B(x,r)$ centered at x with radius r.

In other contexts, the Hardy-Littlewood maximal function is defined differently, by taking the supremum over all balls containing x, rather than those centered at x. However, one can easily see that the two functions in the definitions are comparable.

In a natural way, the Hardy Littlewood maximal function can be defined for a measure μ in the following way:

$$M(\mu)(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} d\mu(y), \tag{95}$$

where μ is a measure on $\mathbb{R}^d$ and $B(x,r)$ is a ball centered at x with radius r.

Remark A.26. Sometimes one works with a *truncated* maximal function, where the supremum in (94) is taken only over balls of radius at most R. In that setting, essentially all of the results continue to hold, and the proofs carry over with only minor adjustments.

Maximal functions arise in several contexts, and it is crucial to understand their properties and the spaces in which they act as bounded operators. In particular, we have the following theorem, whose proof can be found in Chapter 1 of [**Ste93**].

Theorem A.27 The Hardy-Littlewood maximal function M is bounded from $L^p(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$ for $1 < p \leq \infty$, and satisfies the weak-type inequality for $L^1(\mathbb{R}^d)$.

In particular, there exists a constant C_p such that, for any $f \in L^p(\mathbb{R}^d)$, with $1 < p \leq \infty$,

$$\|M(f)\|_{L^p} \leq C_p \|f\|_{L^p}.$$

Additionally, the Hardy-Littlewood maximal function satisfies the weak-type inequality:

$$\mathscr{L}^d \left(\left\{ x \in \mathbb{R}^d : M(f)(x) > \lambda \right\} \right) \leq \frac{C_1}{\lambda} \|f\|_{L^1},$$

for any $f \in L^1(\mathbb{R}^d)$ and for $\lambda > 0$, where C_1 is a constant.

Sometimes we will need interpolation inequalities in order to compensate for the lack of a strong type inequality for $p = 1$. We define the weak Lebesgue space as follows:

$$\|f\|_{L^{1,\infty}} := \sup_{\lambda > 0} \lambda \, \mathscr{L}^d \left(\{ x \in \mathbb{R}^d : |f(x)| > \lambda \} \right). \tag{96}$$

One can see that the maximal function is an operator that sends L^1 into $L^{1,\infty}$. If the function has further properties, being in that space can be improved to stay in L^1.

Proposition A.28 Let f be a function with compact support such that $f \in L^\infty \cap L^{1,\infty}$. Then $f \in L^1$, and the following estimate holds:

$$\|f\|_{L^1} \leq \|f\|_{L^{1,\infty}} \left(1 + \log\left(\frac{C\|f\|_{L^\infty}}{\|f\|_{L^{1,\infty}}}\right)\right),$$

where $C = \mathscr{L}^d(\operatorname{supp} f)$.

Proof. Let $A := \|f\|_{L^\infty}$ and $B := \|f\|_{L^{1,\infty}}$ for clarity. By the layer cake representation formula, we have

$$\|f\|_{L^1} = \int_0^\infty \mathscr{L}^d\left(\left\{x \in \mathbb{R}^d : |f(x)| \geq \lambda\right\}\right) d\lambda = \int_0^A \mathscr{L}^d\left(\left\{x \in \mathbb{R}^d : |f(x)| \geq \lambda\right\}\right) d\lambda.$$

Fix now k a parameter to be chosen later, we can thus rewrite We estimate the integral as follows:

$$\begin{aligned}\int_0^A \mathscr{L}^d\left(\left\{x \in \mathbb{R}^d : |f(x)| \geq \lambda\right\}\right) d\lambda &= \int_0^k \mathscr{L}^d\left(\left\{x \in \mathbb{R}^d : |f(x)| \geq \lambda\right\}\right) d\lambda \\ &\quad + \int_k^A \mathscr{L}^d\left(\left\{x \in \mathbb{R}^d : |f(x)| \geq \lambda\right\}\right) d\lambda \\ &\leq kC + \int_k^A \frac{B}{\lambda}\, d\lambda \\ &= kC + B\log\left(\frac{A}{k}\right),\end{aligned}$$

Finally we optimize in k, by choosing $k = \frac{B}{C}$ to get

$$\|f\|_{L^1} \leq \|f\|_{L^{1,\infty}} \left(1 + \log\left(\frac{C\|f\|_{L^\infty}}{\|f\|_{L^{1,\infty}}}\right)\right),$$

as wanted. □

For convenience, given a Sobolev function f, we adopt the following convention: the function is defined pointwise by the value

$$\lim_{r\to 0} \frac{1}{\omega_d r^d} \int_{B_r(x)} f(y)\, dy$$

when $x \in \mathbb{R}^d$ is a Lebesgue point for f, and zero otherwise.

We now prove the following crucial inequality in the context of Sobolev spaces.

Proposition A.29 Let $f \in W^{1,p}(\mathbb{R}^d)$ for some $1 \leq p \leq \infty$. Then, for almost every $x, y \in \mathbb{R}^d$, it holds that

$$|f(x) - f(y)| \lesssim |x - y| \left(M(|\nabla f|)(x) + M(|\nabla f|)(y)\right).$$

Proof. Fix $x, y \in \mathbb{R}^d$, and define $r := |x - y|$. Consider the set $A_r := B_r(x) \cap B_r(y)$. It is easy to verify that $\mathscr{L}^d(A_r) = K\,\mathscr{L}^d(B_r(x))$, for some constant K depending only on the dimension d. We start from the identity

$$|f(x) - f(y)| = \frac{1}{|A_r|} \int_{A_r} |f(x) - f(y)|\, dz.$$

Using the triangle inequality, we estimate

$$|f(x) - f(y)| \leq \frac{1}{|A_r|} \int_{A_r} |f(x) - f(z)|\, dz + \frac{1}{|A_r|} \int_{A_r} |f(z) - f(y)|\, dz.$$

Remember that $|x - z| \leq r$, so we can deduce that

$$|f(x) - f(y)| \lesssim \frac{r}{|B_r(x)|} \int_{B_r(x)} \frac{|f(x) - f(z)|}{|x - z|}\, dz + \frac{r}{|B_r(y)|} \int_{B_r(y)} \frac{|f(y) - f(z)|}{|y - z|}\, dz.$$

We now use the fundamental theorem of calculus to estimate each term. In particular,

$$\frac{|f(x)-f(z)|}{|x-z|} \leq \int_0^1 |\nabla f(x+s(z-x))|\, ds,$$

and therefore

$$\int_{B_r(x)} \frac{|f(x)-f(z)|}{|x-z|}\, dz \leq \int_{B_r(x)} \int_0^1 |\nabla f(x+s(z-x))|\, ds\, dz.$$

Finally, we perform the change of variables $\tilde{z} = s(z-x)$, and apply Tonelli's Theorem to exchange the order of integration. We obtain:

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} \int_0^1 |\nabla f(x+s(z-x))|\, ds\, dz = \int_0^1 \left(\frac{1}{|B_r(x)|} \int_{B_{sr}(0)} |\nabla f(x+\tilde{z})| \cdot s^{-d}\, d\tilde{z} \right) ds,$$

where we have used the fact that $d\tilde{z} = s^d\, dz$, and the set $\tilde{z} \in s(B_r(x) - x) = B_{sr}(0)$. Thus:

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} \int_0^1 |\nabla f(x+s(z-x))|\, ds\, dz = \int_0^1 \frac{1}{|B_r(x)|} \cdot \frac{1}{s^d} \int_{B_{sr}(x)} |\nabla f(w)|\, dw\, ds.$$

Finally, by using the definition of the maximal function, we obtain

$$r \cdot \frac{1}{|B_r(x)|} \int_{B_r(x)} \int_0^1 |\nabla f(x+s(z-x))|\, ds\, dz \leq r\, M(|\nabla f|)(x).$$

Applying the same argument at the point y, and recalling that $r = |x-y|$, we conclude:

$$|f(x)-f(y)| \lesssim |x-y| \left(M(|\nabla f|)(x) + M(|\nabla f|)(y) \right).$$

□

Remark A.30. We emphasize that the proof is local in nature, so the statement of the proposition still holds if f belongs to a local Sobolev space, i.e., $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^d)$.

A.3. Results about the Littlewood–Paley Decomposition

In this appendix, we prove several results about the Littlewood–Paley decomposition that are used and referenced in Chapter 4. For the sake of exposition, we work with the Littlewood–Paley decomposition of a function on $\mathbb{R}^d$, rather than that of a mean-free function on $\mathbb{T}^d$. The only difference in the proofs lies in the treatment of low frequencies: in the mean-free case, only Littlewood–Paley components with $k \geq 1$ appear, whereas in the general case, the decomposition also includes blocks with indices $k < 0$. As in Chapter 4, we start with a radial Schwartz function φ whose Fourier transform $\widehat{\varphi}$ is supported in the unit ball $B_1(0)$. Moreover, $\widehat{\varphi}$ satisfies

$$0 \leq \widehat{\varphi}(\xi) \leq 1 \quad \text{for all } \xi,$$

and equals 1 on the closed ball $\overline{B_{\frac{1}{2}}(0)}$. This function will be used to construct a dyadic partition of unity in frequency space. For each $k \in \mathbb{Z}$, we define

$$\varphi_k(x) := 2^{kd}\varphi(2^k x) - 2^{(k-1)d}\varphi(2^{k-1}x),$$

so that the Fourier transform $\widehat{\varphi}_k$ is supported in a dyadic annulus:

$$\operatorname{spt} \widehat{\varphi}_k \subset B_{2^k}(0) \setminus \overline{B_{2^{k-2}}(0)}.$$

As a result, each φ_k overlaps in frequency only with its immediate neighbors. Notice that is defined in such a way that

$$\widehat{\varphi}_k(\xi) = \widehat{\varphi}(2^{-k}\xi) - \widehat{\varphi}(2^{-(k-1)}\xi).$$

From this last equation and the definition of φ we deduce that

$$\widehat{\varphi}_k(\xi) = \widehat{\varphi}_k(\xi)(\widehat{\varphi}_{k-1}(\xi) + \widehat{\varphi}_k(\xi) + \widehat{\varphi}_{k+1}(\xi))$$

since $\widehat{\varphi}_{k-1}(\xi) + \widehat{\varphi}_k(\xi) + \widehat{\varphi}_{k+1}(\xi) = 1$ on the support of $\widehat{\varphi}_k(\xi)$. By applying the inverse Fourier transform, we obtain

$$\varphi_k = \varphi_k * (\varphi_{k-1} + \varphi_k + \varphi_{k+1}).$$

Notice also that the family $\{\varphi_k\}_{k\in\mathbb{Z}}$ is constructed to form a partition of unity in frequency space, i.e.,

$$\sum_{k\in\mathbb{Z}} \widehat{\varphi}_k(\xi) = 1 \quad \text{for all } \xi \in \mathbb{R}^d \setminus \{0\}.$$

Furthermore, for integers $n \le m$, we introduce

$$\psi_{n\le m}(x) := \sum_{j=n}^{m} \varphi_j(x).$$

Finally, given a function $f \in L^1_{\text{loc}}$, we define the Littlewood–Paley blocks f_k as

$$f_k := f * \varphi_k.$$

We also define

$$f_{n\le m} := f * \psi_{n\le m},$$

and

$$f_k^{\le} := f * \left(2^{kd}\varphi\left(\frac{\cdot}{2^k}\right)\right),$$

which corresponds, in fact, to the case $n = -\infty$ in the previous definition.

Notice that by applying the Fourier transform, these definitions are equivalent to products in frequency space. More precisely, we have

$$\widehat{f_k}(\xi) = \widehat{f}(\xi)\,\widehat{\varphi_k}(\xi),$$

while for the sum over blocks,

$$\widehat{f_{n\le m}}(\xi) = \widehat{f}(\xi)\,\widehat{\psi_{n\le m}}(\xi) = \widehat{f}(\xi)\sum_{j=n}^{m}\widehat{\varphi_j}(\xi).$$

Finally, for the low-frequency cutoff,

$$\widehat{f_k^{\le}}(\xi) = \widehat{f}(\xi)\,\widehat{\varphi}\left(\frac{\xi}{2^k}\right).$$

Sometimes, we will also use the high-frequency cutoff defined by

$$f_k^{\ge} := f - f_{k-1}^{\le}.$$

By (78), both $f_k(x)$ and $f_k^{\le}(x)$ are bounded, up to a constant, by $Mf(x)$. Hence, they belong to L^p whenever $f \in L^p$ with $p > 1$. We start by proving this proposition that gives a sense in which $f = \sum f_k$ holds.

Proposition A.31 For $1 < p < \infty$ and $f \in L^p$, the sum $\sum_k f_k$ converges to f in L^p.

Proof. Note that $f_{n\le m}(x) = f_m^{\le}(x) - f_{n-1}^{\le}(x)$. Hence, it suffices to show that $f_m^{\le} \to f$ in L^p as $m \to +\infty$, and that $f_n^{\le} \to 0$ in L^p as $n \to -\infty$. We start by proving the first of the two statements(that holds also for $p = 1$) by observing that

$$\|f - f_k^{\le}\|_{L^p} = \left(\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\varphi(y)\,[f(x) - f(x - \tfrac{y}{2^k})]\,dy\right|^p dx\right)^{1/p},$$

where we have used a change of variable to rewrite the convolution, and the fact that $\int\varphi = 1$, which follows from $\widehat{\varphi}(0) = 1$. By applying Minkowski's inequality, we obtain

$$\|f - f_k^{\le}\|_{L^p} \le \int_{\mathbb{R}^d}|\varphi(y)|\cdot\|f(\cdot - \tfrac{y}{2^k}) - f(\cdot)\|_{L^p}\,dy.$$

Fix $\varepsilon > 0$. Since $\varphi \in L^1(\mathbb{R}^d)$, we can choose $R > 0$ such that

$$\int_{B_R(0)^c}|\varphi(y)|\,dy \le \frac{\varepsilon}{2\|f\|_{L^p}}.$$

We split the integral:

$$\|f - f_k^{\le}\|_{L^p} \le \int_{B_R(0)}|\varphi(y)|\,\|f(\cdot - \tfrac{y}{2^k}) - f(\cdot)\|_{L^p}\,dy + \int_{B_R(0)^c}|\varphi(y)|\,\|f(\cdot - \tfrac{y}{2^k}) - f(\cdot)\|_{L^p}\,dy.$$

The second term is bounded by

$$2\|f\|_{L^p}\int_{B_R(0)^c}|\varphi(y)|\,dy\le\varepsilon.$$

On the other hand, for $y\in B_R(0)$, the quantity $\|f(\cdot-\frac{y}{2^k})-f(\cdot)\|_{L^p}\to 0$ as $k\to\infty$, and is dominated by $2\|f\|_{L^p}$. So by dominated convergence,

$$\lim_{k\to\infty}\int_{B_R(0)}|\varphi(y)|\,\|f(\cdot-\tfrac{y}{2^k})-f(\cdot)\|_{L^p}\,dy=0.$$

Hence for k large enough, the first integral is also $\le\varepsilon$, and we conclude that

$$\lim_{k\to\infty}\|f-f_k^{\le}\|_{L^p}=0.$$

We now finish to prove the proposition by proving the second claim. Since $|f_k^{\le}|(x)\lesssim Mf(x)\in L^p$, it suffices, thanks to dominated convergence theorem to show that $f_k^{\le}\to 0$ pointwise. By a simple estimate, if f is a Schwartz function, we have

$$|f_k^{\le}(x)|=\left|\int_{\mathbb{R}^d}f(x-y)\,2^{kd}\varphi\left(\frac{y}{2^k}\right)\,dy\right|\le 2^{kd}\|f\|_{L^1}\|\varphi\|_{L^\infty}.$$

This shows that $f_k^{\le}(x)\to 0$ pointwise thus it follows that $f_k^{\le}\to 0$ in L^p by the previous cited dominated convergence. To conclude the argument for general $f\in L^p$, notice that the map $f\mapsto f_k^{\le}$ is bounded from L^p to L^p uniformly in k, as shown earlier using the Hardy–Littlewood maximal function. Since the Schwartz space is dense in L^p, the convergence result extends to all $f\in L^p$. □

We now state a proposition that links the L^p norm of the frequency blocks with their derivatives, commonly known as the Bernstein inequalities.

Proposition A.32 (Bernstein inequalities) Let $1\le p\le\infty$ and $s\in\mathbb{R}$. Then the following equivalence holds:

$$\||\nabla|^s f_k\|_{L^p}\sim 2^{ks}\|f_k\|_{L^p}.$$

Moreover, for all $s>0$, the following one-sided estimates are valid:

$$\||\nabla|^s f_k^{\le}\|_{L^p}\lesssim 2^{ks}\|f\|_{L^p},$$

$$\|f_k^{\ge}\|_{L^p}\lesssim 2^{-ks}\||\nabla|^s f\|_{L^p}.$$

We recall with a definition what we mean by $|\nabla|^s$.

Definition A.33. For $s\in\mathbb{R}$, the operator $|\nabla|^s$ is defined via the Fourier transform as

$$|\nabla|^s f:=\mathcal{F}^{-1}\left(|\xi|^s\widehat{f}(\xi)\right),$$

where $\widehat{f}$ denotes the Fourier transform of f, and $\mathcal{F}^{-1}$ its inverse.

In particular, when f is sufficiently regular (e.g., $f\in\mathcal{S}(\mathbb{R}^d)$), we have:

$$|\nabla|f(x)=|\nabla f(x)|=\left(\sum_{j=1}^d|\partial_j f(x)|^2\right)^{1/2},\quad |\nabla|^2 f=-\Delta f.$$

Proof. It suffices to prove the estimates for f_k, as the others follow from summation and the triangle inequality in L^p. For example,

$$\||\nabla|^s f_k^{\le}\|_{L^p}\le\sum_{j=-\infty}^{k}\||\nabla|^s f_j\|_{L^p}\lesssim\sum_{j=-\infty}^{k}2^{js}\|f_j\|_{L^p}\lesssim 2^{ks}\|f\|_{L^p},$$

where we also used that the L^p-norm of each block is controlled by the L^p-norm of the function. The estimate for $f_k^{\ge}$ follows by a similar argument. We now turn to the proof of the main estimate. It suffices to prove the inequality

$$\||\nabla|^s f_k\|_{L^p}\lesssim 2^{ks}\|f_k\|_{L^p}$$

for every $s \in \mathbb{R}$, since the reverse bound follows by applying the same inequality to $|\nabla|^{-s} f_k$. We want to use Young's convolution inequality. Observe that

$$|\nabla|^s f_k = [(2\pi|\xi|)^s \widehat{\varphi}_k(\xi)]^\vee * f = 2^{ks} \left[\left(2\pi \frac{|\xi|}{2^k} \right)^s \widehat{\psi}_1 \left(\frac{\xi}{2^k} \right) \right]^\vee * f.$$

Define

$$\chi(\xi) := (2\pi|\xi|)^s \widehat{\psi}_1(\xi) \in C_c^\infty(\mathbb{R}^d \setminus \{0\}), \quad \text{and} \quad \chi_k(\xi) := \chi\left(\frac{\xi}{2^k}\right),$$

so that

$$|\nabla|^s f_k = 2^{ks} [\chi_k]^\vee * f = 2^{ks} \left[2^{kd} \chi^\vee(2^k \cdot) \right] * f.$$

Thus, by Young's inequality,

$$\| |\nabla|^s f_k \|_{L^p} \le 2^{ks} \left\| 2^{kd} \chi^\vee(2^k \cdot) \right\|_{L^1} \|f\|_{L^p}.$$

Now notice that

$$\left\| 2^{kd} \chi^\vee(2^k \cdot) \right\|_{L^1} = \|\chi^\vee\|_{L^1},$$

since the scaling preserves the L^1 norm. Hence,

$$\| |\nabla|^s f_k \|_{L^p} \lesssim 2^{ks} \|f\|_{L^p}.$$

To recover the f_k on the right hand side, one can perform the following trick. Instead of using the dyadic partition (in the frequency space) generated by $\widehat{\varphi_1}$—with $\widehat{\varphi_k}(\xi) = \widehat{\varphi_1}(\xi/2^k)$—one considers as a projection family the one generated by $\widehat{\varphi_0} + \widehat{\varphi_1} + \widehat{\varphi_2}$, and denotes the corresponding projection by $\tilde{f}_k$. By repeating the same argument as before, one proves that

$$\| |\nabla|^s \tilde{f}_k \|_{L^p} \lesssim 2^{ks} \|f\|_{L^p},$$

and then, by choosing f_k as f, the proof is complete, since this estimate together with (74) yields the desired result. □

We now focus on proving one of the most important properties of the Littlewood–Paley decomposition. Given a decomposition f_k of a function f, we define the square function associated to f as

$$S(f) := \left(\sum_k |f_k|^2 \right)^{1/2}.$$

Theorem A.34 For every $1 < p < \infty$, we have

$$\|S(f)\|_{L^p} \sim \|f\|_{L^p}.$$

Remark A.35. Note that our proof will also cover (76). One simply needs to follow the same argument of the first part and using the fact that if a function has zero mean, then its Fourier coefficient at frequency zero vanishes.

A crucial tool for making square-function like quantities is the so-called *Khintchine inequality.*

Proposition A.36 (Khintchine inequality) Let $(X_n)_{n \ge 1}$ be a sequence of independent, identically distributed random variables on a probability space, where each X_n takes values ± 1 with equal probability. For any $0 < p < \infty$, and for every sequence $\{c_n\} \in \ell^2$,

$$\left(\mathbb{E} \left| \sum_n c_n X_n \right|^p \right)^{\frac{1}{p}} \sim_p \left(\sum_n |c_n|^2 \right)^{\frac{1}{2}}$$

PROOF. We can assume that $c_n \in \mathbb{R}$. We begin by showing that the two quantities coincide when $p = 2$. In the following computation, we use that $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$, and that the variables X_i are independent:

$$\begin{aligned}\mathbb{E}\left[\left(\sum_i c_i X_i\right)^2\right] &= \mathbb{E}\left[\sum_{i,j} c_i c_j X_i X_j\right] \\ &= \sum_{i,j} c_i c_j \mathbb{E}[X_i X_j] \\ &= \sum_i c_i^2 \mathbb{E}[X_i^2] + \sum_{i \neq j} c_i c_j \mathbb{E}[X_i]\mathbb{E}[X_j] \\ &= \sum_i c_i^2.\end{aligned}$$

We start now by establishing the $\lesssim$ inequality. We estimate the probability of a super-level set using the exponential version of Tchebychev's inequality:

$$\begin{aligned}\mathbb{P}\left(\sum_n c_n X_n > \lambda\right) &\leq e^{-\lambda t}\, \mathbb{E}\left[e^{t \sum_n c_n X_n}\right] \\ &= e^{-\lambda t} \prod_n \mathbb{E}\left[e^{t c_n X_n}\right] \\ &\leq e^{-\lambda t} \prod_n \frac{1}{2}\left(e^{t c_n} + e^{-t c_n}\right) \\ &\leq e^{-\lambda t} \prod_n e^{t^2 c_n^2/2} \\ &= e^{-\lambda t} \cdot e^{t^2 \sum_n c_n^2/2} \\ &= \exp\left(-\lambda t + \frac{t^2}{2}\sum_n c_n^2\right).\end{aligned}$$

Choosing $t = \frac{\lambda}{\sum_n c_n^2}$, we obtain

$$\mathbb{P}\left(\sum_n c_n X_n > \lambda\right) \leq \exp\left(-\frac{\lambda^2}{2\sum_n c_n^2}\right).$$

By symmetry of the random variables X_n, we conclude

$$\mathbb{P}\left(\left|\sum_n c_n X_n\right| > \lambda\right) \leq 2\exp\left(-\frac{\lambda^2}{2\sum_n c_n^2}\right).$$

We proceed by using the layer cake formula:

$$\begin{aligned}\left(\mathbb{E}\left|\sum_n c_n X_n\right|^p\right)^{1/p} &= \left(\int_0^\infty p\lambda^{p-1}\, \mathbb{P}\left(\left|\sum_n c_n X_n\right| > \lambda\right) d\lambda\right)^{1/p} \\ &\leq \left(\int_0^\infty p\lambda^{p-1} \cdot 2\exp\left(-\frac{\lambda^2}{2\sum_n c_n^2}\right) d\lambda\right)^{1/p}.\end{aligned}$$

After the change of variable

$$\mu = \frac{\lambda}{\sqrt{\sum_n c_n^2}} \quad \text{so that} \quad d\lambda = \sqrt{\sum_n c_n^2}\, d\mu,$$

the estimate becomes

$$\begin{aligned}\left(\mathbb{E}\left|\sum_n c_n X_n\right|^p\right)^{1/p} &\leq \left(\int_0^\infty p\lambda^{p-1}\cdot 2\exp\left(-\frac{\lambda^2}{2\sum_n c_n^2}\right)\,d\lambda\right)^{1/p}\\ &= \left(\int_0^\infty p\left(\mu\sqrt{\sum_n c_n^2}\right)^{p-1}\cdot 2\exp\left(-\frac{\mu^2}{2}\right)\cdot\sqrt{\sum_n c_n^2}\,d\mu\right)^{1/p}\\ &\lesssim \left(\sum_n c_n^2\right)^{1/2}\end{aligned}$$

where in the last step we used that the integral in μ was finite. We are left with proving the $\gtrsim$ part of the inequality. We argue first for $p > 1$ by applying Holder's inequality, the equality already established for $p = 2$ and the $\lesssim$ estimate:

$$\begin{aligned}\sum_n c_n^2 &= \mathbb{E}\left[\left|\sum_n c_n X_n\right|^2\right]\\ &\leq \left(\mathbb{E}\left[\left|\sum_n c_n X_n\right|^p\right]\right)^{1/p}\left(\mathbb{E}\left[\left|\sum_n c_n X_n\right|^{p'}\right]\right)^{1/p'}\\ &\leq \left(\mathbb{E}\left[\left|\sum_n c_n X_n\right|^p\right]\right)^{1/p}\left(\sum_n c_n^2\right)^{1/2}.\end{aligned}$$

This proves the Kintchine inequality for $p > 1$, finally we prove the estimate also for $0 < p \leq 1$. The idea is to argue in a similar way

$$\begin{aligned}\sum_n c_n^2 &= \mathbb{E}\left[\left|\sum_n c_n X_n\right|^2\right]\\ &= \mathbb{E}\left[\left|\sum_n c_n X_n\right|^{2-\frac{p}{2}}\cdot\left|\sum_n c_n X_n\right|^{\frac{p}{2}}\right]\\ &\leq \left(\mathbb{E}\left[\left|\sum_n c_n X_n\right|^{4-p}\right]\right)^{1/2}\left(\mathbb{E}\left[\left|\sum_n c_n X_n\right|^p\right]\right)^{1/2}\\ &\lesssim (\sum c_n^2)^{1-p/4}\left(\mathbb{E}\left[\left|\sum_n c_n X_n\right|^p\right]\right)^{1/2},\end{aligned}$$

and this conclude the proof. □

With this proposition we now prove Theorem A.34.

Proof. Let $(X_k)_{k\in\mathbb{Z}}$ be independent, identically distributed random variables with $X_k = \pm 1$ with equal probability. By Khintchine inequality, we have the equivalence

$$S(f)(x) \sim_p \left(\mathbb{E}\left|\sum_k X_k f_k(x)\right|^p\right)^{\frac{1}{p}},$$

thus we have

$$\|S(f)\|_{L^p} \sim_p \mathbb{E}\left[\int\left|\sum_k X_k f_k(x)\right|^p dx\right] = \mathbb{E}\left[\left\|\sum_k X_k f_k\right\|_{L^p}^p\right].$$

The crucial observation is that

$$\sum_k X_k f_k = m_X * f,$$

where the Fourier transform of m_X is given by

$$\widehat{m_X} = \sum_k X_k \widehat{\varphi_k}.$$

In order to prove that $\|S(f)\|_{L^p} \lesssim \|f\|_{L^p}$, it suffices to show that $\widehat{m_X}$ is a Mihlin multiplier with bounds independent of the values of the X_k. Then, the conclusion follows from the Hörmander–Mihlin multiplier theorem; for a detailed proof, see [**Gra14**, Theorem 6.2.7]. The Mihlin condition is easy to verify since

$$|\partial_\xi^\alpha \widehat{m_X}(\xi)| = \left| \sum_k X_k 2^{-k|\alpha|} \partial^\alpha \widehat{\varphi_1}\left(\frac{\xi}{2^k}\right) \right| \lesssim |\xi|^{-|\alpha|},$$

where the last inequality essentially follows from the fact that $\widehat{\varphi_1} \in C_c^\infty(\mathbb{R}^d \setminus \{0\})$, thus only finite terms contribute to the sum. To prove the other inequality, we will exploit the structure of the dyadic partition. Recall from the previous argument the slightly different projection $\tilde{f}_k$; to make the notation more clear, we denote the corresponding operator by $\tilde{P}_k$, while the classical projection $f \mapsto f_k$ will be denoted by P_k. By (74) one can see that $\tilde{P}_k P_k = P_k$. We use a duality argument to show the desired estimate. Let $f \in L^p$, $g \in L^{p'}$. We expand them using the Littlewood–Paley decomposition, thanks to (A.31):

$$\langle f, g \rangle = \int_{\mathbb{R}^d} f(x)\, \bar{g}(x)\, dx = \int_{\mathbb{R}^d} \sum_k f_k(x) \sum_j \bar{g}_j(x)\, dx.$$

By Parseval's identity and the fact that $\tilde{P}_k P_k = P_k$, we see that

$$\begin{aligned}
\langle f, g \rangle &= \int_{\mathbb{R}^d} \sum_k f_k\, \tilde{g}_k\, dx \\
&\leq \int_{\mathbb{R}^d} \left(\sum_k |f_k|^2 \right)^{1/2} \left(\sum_k |\tilde{g}_k|^2 \right)^{1/2} dx \\
&\leq \|S(f)\|_{L^p} \|\widetilde{S}(g)\|_{L^{p'}} \\
&\lesssim \|S(f)\|_{L^p} \|g\|_{L^{p'}}.
\end{aligned}$$

which, by the arbitrariness of $g \in L^{p'}$, gives

$$\|f\|_{L^p} \lesssim \|S(f)\|_{L^p},$$

as wanted. □

We now provide a square-function characterization with gradient blocks.

Proposition A.37 For every $s \in \mathbb{R}$ and $1 < p < \infty$, the following equivalences hold:

$$\||\nabla|^s f\|_{L^p} \sim_{s,p} \left\| \left(\sum_k 2^{2ks} |f_k(x)|^2 \right)^{\frac{1}{2}} \right\|_{L^p},$$

and for $s > 0$

$$\||\nabla|^s f\|_{L^p} \sim_{s,p} \left\| \left(\sum_k 2^{2ks} |f_k^{\geq}(x)|^2 \right)^{\frac{1}{2}} \right\|_{L^p}.$$

Proof. By running the exact same argument as in the first part of the square function estimate, one can prove that

$$\left\| \left(\sum_k 2^{2ks} \left(P_k(|\nabla|^{-s} g) \right)^2 \right)^{1/2} \right\|_{L^p} \lesssim \|g\|_{L^p}.$$

Choosing $g = |\nabla|^s f$, the "$\lesssim$" part of the first equivalence follows. As for the square function estimate, we will prove the inequality $\gtrsim$ by duality. By arguing similar to before, by using Parseval's identity, proposition (A.31), the definition of $|\nabla|^s$, Cauchy Schwartz, Holder inequality and the previous estimate we have

$$\begin{aligned}
\langle g, h\rangle = \int_{\mathbb{R}^d} f\overline{h} &= \int_{\mathbb{R}^d} 2^{ks}|\nabla|^{-s}P_N g\cdot 2^{-2ks}|\nabla|^s\widetilde{P}_N h\,dx \\
&\leq \int_{\mathbb{R}^d}\left(\sum_k 2^{2sk}|P_k|\nabla|^{-s}g|^2\right)^{\frac{1}{2}}\left(\sum_k 2^{-2sk}|\widetilde{P}_k|\nabla|^s h|^2\right)^{\frac{1}{2}} dx \\
&\leq \left\|\left(\sum_k 2^{2sk}|P_k|\nabla|^{-s}g|^2\right)^{\frac{1}{2}}\right\|_{L^p}\left\|\left(\sum_k 2^{-2sk}|\widetilde{P}_k|\nabla|^s h|^2\right)^{\frac{1}{2}}\right\|_{L^{p'}} \\
&\lesssim \left\|\left(\sum_k 2^{2sk}|P_k|\nabla|^{-s}g|^2\right)^{\frac{1}{2}}\right\|_{L^p}\|h\|_{L^{p'}}
\end{aligned}$$

that gives the other desired inequality by duality and by choosing $g = |\nabla|^s f$. Finally, we prove the second inequality by showing that its right-hand side is comparable to the right-hand side of the first inequality in the proposition. It is easy to see, using the identity $f_k = f_k^{\geq} - f_{k+1}^{\geq}$, the inequality $(a-b)^2 \leq 2(a^2+b^2)$, and the monotonicity of the functions involved, that the right-hand side of the second inequality is $\gtrsim$ the right-hand side of the first, indeed, we have

$$\sum_k 2^{2ks}f_k^2 = \sum_k 2^{2ks}\left(f_k^{\geq} - f_{k+1}^{\geq}\right)^2 \leq 2\sum_k 2^{2ks}\left((f_k^{\geq})^2 + (f_{k+1}^{\geq})^2\right) \lesssim \sum_k 2^{2ks}(f_k^{\geq})^2.$$

We turn on the other part, since we are computing an L^p norm we can expand f in Littlewood-Paley blocks, in particular notice that

$$\begin{aligned}
\sum_k 2^{ks}|f_k^{\geq}|^2 &\leq \sum_k 2^{2ks}\sum_{n,m\geq k}|f_n||f_m| \\
&\leq 2\sum_k 2^{2ks}\sum_{n\geq m\geq k}\frac{1}{2^{ns+ms}}\cdot 2^{ns}|f_n|\cdot 2^{ms}|f_m| \\
&= 2\sum_{n\geq m}\sum_{k\leq m}\frac{2^{2ks}}{2^{ns+ms}}\cdot 2^{ns}|f_n|\cdot 2^{ms}|f_m| \\
&\lesssim \sum_{m\leq n}\frac{2^{ms}}{2^{ns}}\cdot 2^{ns}|f_n|\cdot 2^{ms}|f_m| \\
&\lesssim \sum_k 2^{2ks}f_k^2.
\end{aligned}$$

where in the last line we used Schur's test; see Proposition A.38 below. □

Schur's test is a tool used to handle sums of the form

$$\sum_{k,\ell} a_{k\ell}x_k y_\ell$$

when $x, y \in \ell^2$. This fact can be generalized to ℓ^p but we will present just the version we need.

Proposition A.38 Let $a_{k\ell} \geq 0$ be weights such that

$$\sup_k \sum_\ell a_{k\ell} \leq C \quad \text{and} \quad \sup_\ell \sum_k a_{k\ell} \leq C.$$

Then for any sequences $x = (x_k) \in \ell^2$ and $y = (y_\ell) \in \ell^2$, we have

$$\sum_{k,\ell} a_{k\ell}x_k y_\ell \leq C\|x\|_{\ell^2}\|y\|_{\ell^2}.$$

Proof. We can assume x_i, y_j all non-negative. With the use of Cauchy Schwartz two times and some algebraic manipulations we have that

$$\begin{aligned}\sum_{k,\ell} a_{k\ell}x_k y_\ell &= \sum_\ell y_\ell \sum_k a_{k\ell}x_k \le \sum_\ell y_\ell \left(\sum_k a_{k\ell}x_k^2\right)^{1/2}\left(\sum_k a_{k\ell}\right)^{1/2} \\ &\le \left(\sum_\ell y_\ell^2\right)^{1/2}\left(\sum_\ell\sum_k a_{k\ell}x_k^2\cdot\sum_k a_{k\ell}\right)^{1/2} \\ &= \left(\sum_\ell y_\ell^2\right)^{1/2}\left(\sum_k x_k^2\sum_\ell a_{k\ell}\sum_j a_{j\ell}\right)^{1/2} \\ &\le \|y\|_{\ell^2}\left(\sup_\ell\sum_j a_{j\ell}\cdot\sum_k x_k^2\sum_\ell a_{k\ell}\right)^{1/2} \\ &\le \|x\|_{\ell^2}\|y\|_{\ell^2}\cdot\left(\sup_k\sum_\ell a_{k\ell}\cdot\sup_\ell\sum_k a_{k\ell}\right)^{1/2} \\ &\le C\|x\|_{\ell^2}\|y\|_{\ell^2},\end{aligned}$$

as wanted. □

Remark A.39. In the sum

$$\sum_{m\le n}\frac{2^{ms}}{2^{ns}}\cdot 2^{ns}|f_n|\cdot 2^{ms}|f_m|,$$

the weights were given by $a_{m,n} = 1_{m\le n}\frac{2^{ms}}{2^{ns}}$, while the sequences are $x_n = 2^{ns}|f_n|$ and $y_m = 2^{ms}|f_m|$.

Note also the following thing, by proposition A.37, we also have that

$$\|f\|_{L^p} \sim \left\|\left(\sum_k 2^{-2k}|\nabla f_k|^2\right)^{1/2}\right\|_{L^p}. \tag{97}$$

We now want to prove a generalized square function estimate involving gradients, in the case where we have a weighted square function

$$\left(\sum_k\left|\frac{f_k}{\omega(2^{-k})}\right|^2\right)^{1/2}$$

in L^p where ω is an Osgood modulus of continuity. Consider now the function

$$g = \sum_k \frac{f_k}{\omega(2^{-k})},$$

and note that, by Theorem 75, it is bounded in L^p since its Littlewood–Paley blocks carry weights comparable to $\omega(2^{-k})$. Hence, g is finite almost everywhere and belongs to L^p. By applying (97) to g, we obtain

$$\left\|\left(\sum_k\left|\frac{f_k}{\omega(2^{-k})}\right|^2\right)^{1/2}\right\|_{L^p} \sim \left\|\left(\sum_k\left|\frac{2^{-k}\nabla f_k}{\omega(2^{-k})}\right|^2\right)^{1/2}\right\|_{L^p}, \tag{98}$$

and by iterating the same reasoning, we also get

$$\left\|\left(\sum_k\left|\frac{f_k}{\omega(2^{-k})}\right|^2\right)^{1/2}\right\|_{L^p} \sim \left\|\left(\sum_k\left|\frac{2^{-2k}\nabla^2 f_k}{\omega(2^{-k})}\right|^2\right)^{1/2}\right\|_{L^p}. \tag{99}$$

Bibliography

[ABL04] Luigi Ambrosio, François Bouchut, and Camillo De Lellis, *Well-posedness for a class of hyperbolic systems of conservation laws in several space dimensions*, Communications in Partial Differential Equations **29** (2004), no. 10–12, 1635–1651.

[AC14] L. Ambrosio and G. Crippa, *Continuity equations and ode flows with non-smooth velocity*, Proceedings of the Royal Society of Edinburgh **144A** (2014), 1191–1244.

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, Oxford University Press, 2000.

[Amb04] Luigi Ambrosio, *Transport equation and cauchy problem for bv vector fields*, Inventiones mathematicae **158** (2004), no. 2, 227–260.

[BC13] François Bouchut and Gianluca Crippa, *Lagrangian flows for vector fields with gradient given by a singular integral*, Journal of Hyperbolic Differential Equations **10** (2013), no. 2, 235–282.

[Bil99] Patrick Billingsley, *Convergence of probability measures*, 2nd ed., Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, 1999, A Wiley-Interscience Publication.

[BN20] Elia Bruè and Quoc-Hung Nguyen, *On the sobolev space of functions with derivative of logarithmic order*, Advances in Nonlinear Analysis **9** (2020), no. 1, 836–849.

[BN21] ———, *Sharp regularity estimates for solutions of the continuity equation drifted by sobolev vector fields*, Analysis & PDE **14** (2021), no. 8, 2539–2559.

[CL08] Gianluca Crippa and Camillo De Lellis, *Estimates and regularity results for the diperna-lions flow*, Journal für die Reine und Angewandte Mathematik (Crelle's Journal) **616** (2008), 15–46.

[Cri07] Gianluca Crippa, *The flow associated to weakly differentiable vector fields*, Ph.D. thesis, Scuola Normale Superiore di Pisa; Università di Zurigo, 2007.

[DL89] R. J. DiPerna and P.-L. Lions, *Ordinary differential equations, transport theory and sobolev spaces*, Inventiones Mathematicae **98** (1989), no. 3, 511–547.

[EG92] Lawrence C. Evans and Ronald F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[Gra14] Loukas Grafakos, *Classical fourier analysis*, 3rd ed., Graduate Texts in Mathematics, vol. 249, Springer, 2014.

[Gut22] Larry Guth, *Decoupling estimates in fourier analysis*, Proceedings of the International Congress of Mathematicians 2022, vol. 2, EMS Press, 2022, pp. 1054–1089.

[Ha96] Piotr Hajł asz, *Sobolev spaces on an arbitrary metric space*, Potential Analysis **5** (1996), no. 4, 403–415.

[LL15] Huaiqian Li and Dejun Luo, *A unified treatment for odes under osgood and sobolev type conditions*, Bulletin des Sciences Mathématiques **139** (2015), no. 1, 114–133.

[MS24] David Meyer and Christian Seis, *Propagation of regularity for transport equations. a littlewood–paley approach*, Indiana University Mathematics Journal **73** (2024), no. 2, 445–473.

[Ste93] Elias M. Stein, *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*, PMS-43, Princeton University Press, Princeton, NJ, 1993.