

Maximal mixing by incompressible fluid flows

Christian Seis*

September 2, 2013

Abstract

We consider a model for mixing binary viscous fluids under an incompressible flow. We prove the impossibility of perfect mixing in finite time for flows with finite viscous dissipation. As measures of mixedness we consider a Monge–Kantorovich–Rubinstein transportation distance and, more classically, the H^{-1} norm. We derive rigorous a priori lower bounds on these mixing norms which show that mixing cannot proceed faster than exponentially in time. The rate of the exponential decay is uniform in the initial data.

1 Introduction

The present manuscript is concerned with optimal stirring strategies for binary mixtures of incompressible viscous fluids. More precisely, we study decay rates of certain mixing norms with respect to a constrained velocity field. We focus on passive scalar mixing, which means that the feedback of the transported quantity on the flow field is negligible. To model the binary mixture, we consider an indicator function $\rho = \rho(t, x)$ which takes the values $+1$ and -1 only, so that the sets $\{\rho = 1\}$ and $\{\rho = -1\}$ represent the regions in which the fluid consists of component ‘A’ and component ‘B’, respectively. As usual, t and x are the time and space variable, respectively. The stirring velocity field will be denoted by $u = u(t, x)$, and we assume this vector field to be smooth. The transport of the passive scalar by the incompressible flow is then described by the system

$$\partial_t \rho + u \cdot \nabla \rho = 0, \tag{1}$$

$$\nabla \cdot u = 0, \tag{2}$$

and we impose the initial condition $\rho(0, x) = \rho_0(x) \in \{\pm 1\}$. For mathematical convenience, we finally assume that all quantities are periodic in the spatial variables with period cell $[0, 1)^d$. Observe that the total mass of each species is preserved under

*Department of Mathematics, University of Toronto, 40 St. George Street, M5S 2E4, Toronto, Ontario, Canada

the flow. In the case of a critical mixture, that is the case where both components have the same volume fraction and which we assume in this paper, this means

$$\int \rho(t, \cdot) dx = \int \rho_0 dx = 0 \quad \text{for all } t > 0.$$

A natural question, first formulated in [9], addresses the optimality of stirring strategies: How efficient can a binary fluid be mixed by an incompressible flow? Classical measures of the mathematical quantification of the degree of mixedness are negative Sobolev norms, with the homogeneous H^{-1} norm leading the way — not only because of its analytical accessibility. While any \dot{H}^{-s} norm with $s > 0$ can be considered as a mixing norm, the \dot{H}^{-1} is distinguished as it scales like a length, and thus, the actual value of this norm relates to the typical size of one-component (i.e., unmixed) regions in the mixture. Negative Sobolev norms measure oscillations of the indicator function: The larger the typical length scales, the larger the negative Sobolev norm. The actual value of s corresponds to a weight in the measurement of the length scales. On the upper end, the H^0 (i.e., L^2) norm is the variance, which is conserved under the flow (1)&(2), and reveals thus no information on the domain morphology. In the context of mixing, negative Sobolev norms were used as large-scale mixing measures in [13, 6, 17, 9]. Alternative measures of the degree of mixedness can be borrowed from the theory of optimal transportation: Monge–Kantorovich–Rubinstein (MKR) distances can be used to measure the “distance” between two configurations, for instance, the current state and the perfectly mixed state. These MKR distances have been successfully used as measures of mixedness in related studies on demixing problems [8, 15, 18, 2, 16]. We will review the definition of the MKR mixing measure in Subsection 2 below. For more notions of mixing measures, we refer to [9, 11] and references therein.

Any quantitative statement about the efficiency of stirring strategies necessitates to specify certain constraints on the flow, for instance, by fixing the budget for the kinetic energy $\|u\|_{L^2}$, the viscous dissipation $\|\nabla u\|_{L^2}$, or the palenstrophy $\|\Delta u\|_{L^2}$. Under such constraints, several optimal control problems were proposed to study effective mixing, e.g., [12, 10, 7]. In practice, however, solving optimal control problems may be difficult, so different approaches that are suitable also for numerical studies are desirable. The authors in [9, 11] seek to find optimal stirring strategies by choosing the velocity field u in each time step in direction of the steepest descent of the mixing norms. Their strategy also includes identifying absolute limits on how fast perfect mixing could ever be achieved subject to given constraints on the stirring flow. Here, the choice of the constraint is crucial: While optimal stirring strategies with a fixed energy budget $\|u\|_{L^2}$ yield to perfect mixing in finite time [11, Sec. 2], the \dot{H}^{-1} mixing norm cannot decay faster than exponentially when stirring under a palenstrophy constraint $\|\Delta u\|_{L^2}$, cf. [11, Sec. 3]. Such statements on the limit of stirring strategies result from rigorous *a priori* lower bounds on the mixing norm. In the present work, we address mixing with a prescribed viscous dissipation $\|\nabla u\|_{L^2}$ (or even more general $\|\nabla u\|_{L^p}$ with $p \in (1, \infty]$). This case is most natural

for engineering applications where the focus is on overcoming viscous dissipation to maintain stirring. We derive rigorous lower bounds on both MKR mixing measures and \dot{H}^{-1} norms showing that these cannot decay faster than exponentially in time. The bounds are in agreement with the exponential rates observed during direct numerical simulations of optimal stirring, as reported in [9].

The remainder of this paper is organized as follows: In Section 2 we introduce the MKR mixing measure and we present our main results. Section 3 contains the proofs. In the appendix, we finally discuss an important tool in our analysis, namely weighted local maximal functions.

2 Mixing norms and main results

Our primary result, the impossibility of perfect mixing in finite time, comes as a lower bound on an MKR mixing measure. Before stating the theorem, we provide the basic notation. Since MKR distances are the subject of the theory of optimal transportation, which is itself a wide and rapidly growing area of research, we will be unobtrusive in compiling general facts on MKR distances and refer the interested reader to the excellent monographs [20, 21].

Definition 1 (Monge–Kantorovich–Rubinstein distance). *Let $\rho_0, \rho_1 : [0, 1]^d \rightarrow [0, \infty)$ be two periodic probability densities, i.e.,*

$$\int \rho_0 dx = 1 = \int \rho_1 dy.$$

We denote by $\Gamma(\rho_0, \rho_1)$ the set of all probability measures on the product space $[0, 1]^d \times [0, 1]^d$ with marginals ρ_0 and ρ_1 , i.e.,

$$\iint \zeta(x) d\pi(x, y) = \int \zeta \rho_0 dx, \quad \iint \zeta(y) d\pi(x, y) = \int \zeta \rho_1 dy,$$

for all periodic continuous functions ζ on $[0, 1]^d$. Let $\sigma : [0, \infty) \rightarrow \mathbb{R} \cup \{-\infty\}$ denote a continuous function. The Monge–Kantorovich–Rubinstein (MKR) distance with cost function σ is defined as

$$d_c(\rho_0, \rho_1) := \inf_{\pi \in \Gamma(\rho_0, \rho_1)} \iint \sigma(|x - y|) d\pi(x, y).$$

Roughly speaking, the MKR distance $d_\sigma(\rho_0, \rho_1)$ measures the minimal total cost that is required for transferring mass from a source (producer) ρ_0 to a sink (consumer) ρ_1 . The minimal total cost depends on the cost function $\sigma(z)$ which encodes the price one has to pay when transferring a unit of mass over a distance z . The measures $\pi \in \Gamma(\rho_0, \rho_1)$ are the transport plans, i.e., $\pi(x, y)$ is the amount of mass that is transported from x to y .

With the help of the MKR distance, we introduce the following mixing measure:

Definition 2 (MKR mixing measure). *Let $\rho : [0, 1]^d \rightarrow \{\pm 1\}$ be a periodic function with*

$$\int \rho dx = 0.$$

Let ρ_+ and ρ_- be the probability densities defined by

$$\rho_+ = 2 \max\{\rho, 0\} \quad \text{and} \quad \rho_- = -2 \min\{\rho, 0\},$$

so that $\rho = \frac{1}{2}(\rho_+ - \rho_-)$. We then define the MKR mixing measure as

$$D(\rho) := \exp(d_{\ln}(\rho_+, \rho_-)).$$

Notice that even though the above definition can be easily extended to the vector space of mean-zero functions on $[0, 1]^d$, the quantity $D(\rho)$ is not a norm as it lacks the triangle inequality.

We remark that the above MKR mixing measure has units of length and describes thus, at least heuristically, the typical size of the unmixed regions in the mixture. Moreover, by the virtue of being defined via an MKR distance, $D(\rho)$ measures the total cost required for transporting ρ_+ into ρ_- , which is decreasing when the typical size of the unmixed regions gets smaller, that is, when the fluid is better mixed.

Similar logarithmic MKR distances were used in [2, 16] in the study of demixing problems, where the flow field u is slaved to the indicator function ρ via a Stokes equation, modelling spinodal decomposition in binary viscous fluids. In a certain sense, our first result comes as a corollary of a dissipation inequality derived in [2, Proposition 2.2].

Theorem 1 ([2]). *Let $1 < p \leq \infty$. There exists a constant $c > 0$ depending on p and d only such that for every $T > 0$*

$$D(\rho(T)) \geq D(\rho_0) \exp\left(-c \int_0^T \|\nabla u\|_{L^p} dt\right).$$

The statement in Theorem 1 shows that the MKR mixing measure cannot decay faster than exponentially in time, and thus, perfect mixing in finite time is impossible for *any* given velocity vector field u for which

$$\int_0^T \|\nabla u\|_{L^p} dt < \infty \quad \text{for all } T > 0.$$

Remark 1 (Possible generalizations). *We remark there is room for several generalizations of the above result:*

1. *An analogous bound can be derived in the case where $m := \int \rho dx \in (-1, 1) \setminus \{0\}$. In this situation, the definition of $D(\rho)$ has to be modified: ρ_{\pm} has to be replaced by $(\rho - m)_{\pm}$, where $(\cdot)_{\pm}$ is defined as in Definition 2.*

2. As the system (1)&(2) is invariant under the rescaling $x = L^{-1}\hat{x}$, $t = L^{-1}\hat{t}$, $\rho = \hat{\rho}$, and $u = \hat{u}$, our results extend to any domain $[0, L]^d$ with $L > 0$. In this case, one has to understand the involved spatial integrals as spatial averages.
3. The assumption that $\rho \in \{\pm 1\}$ is rather for notational convenience. Qualitatively similar results are true for any smooth density $\rho \in \mathbb{R}$ with $\|\rho\|_{L^{p'}} \leq 1$ where $1/p + 1/p' = 1$.

In any of these case, the constants appearing in the statement of Theorem 1 remain generic, i.e., they depend on d and p only.

Notice that the statement in Theorem 1 is closely related to a mixing conjecture of Bressan [3]. A version of this conjecture has been already established in [5] and our approach uses techniques developed therein. The connection between Bressan's conjecture and \dot{H}^{-1} decay rates is discussed in [11, Sec. 4.A].

Although the MKR mixing measure is not known to be equivalent to a (negative) Sobolev norm, the following estimates, derived in Lemmas 1 and 2 below, indicate that MKR distances yield indeed reasonable measures of the degree of mixedness:

$$\frac{c}{[\rho]_{BV}} \leq D(\rho) \leq \|\rho\|_{\dot{H}^{-1}}. \quad (3)$$

Here, $c > 0$ is a uniform constant that depends only on the space dimension and $[\rho]_{BV}$ is the BV norm of the indicator function ρ . If the boundary of the level sets of ρ are sufficiently regular, we simply have $[\rho]_{BV} = 2\mathcal{H}^{d-1}(\{\rho = 1\})$.

As a corollary of Theorem 1 and (3), we obtain the following estimate on the decay rate of the \dot{H}^{-1} norm.

Theorem 2. *Let $1 < p \leq \infty$. There exists constants $c, C > 0$ depending on p and d only such that for every $T > 0$*

$$[\rho_0]_{BV} \|\rho(T, \cdot)\|_{\dot{H}^{-1}} \geq C \exp\left(-c \int_0^T \|\nabla u\|_{L^p} dt\right).$$

This second statement excludes the possibility of finite time mixing in terms of the more classical \dot{H}^{-1} mixing norm. The lower bound is in agreement with the robust numerical results obtained in [9], where exponential decay rates were observed for the \dot{H}^{-1} mixing norm when applying optimal stirring strategies. Unfortunately, our approach does not allow for a measurement of the degree of mixedness of the initial configuration in terms of this \dot{H}^{-1} norm, but only in terms of the inverse BV norm. The latter norm is weaker in the sense that $[\rho]_{BV}^{-1} \lesssim \|\rho\|_{\dot{H}^{-1}}$, see (3). Notice that (inverse) BV norms are suitable measures of demixing problems [4, 2] rather than mixing problems, as small fluctuations along the boundary of $\{\rho = 1\}$ increase the perimeter while the typical length scale of the unmixed regions (and thus the \dot{H}^{-1} norm) remains unchanged. Hence, smallness of $[\rho]_{BV}^{-1}$ does not imply that the typical length scale is small. For sufficiently regular initial configuration (e.g. stripe

or checkerboard pattern), however, both quantities are morally equivalent. The interplay of BV and \dot{H}^{-1} norms has been the subject of intensive research in recent years, see e.g. [1].

Remark 2 (Possible generalizations). *Again, we discuss possible generalizations of the statement of Theorem 2.*

1. *With the same modifications as in Remark 1, the analogous result applies when considering arbitrary systems $[0, L]^d$ with $L > 0$.*
2. *In the case of off-critical mixtures, e.g., $1 > m \gg 0$, an analogous result is not obvious. Our proof of estimating the MKR mixing norm by the \dot{H}^{-1} norm is quite crude as it relies essentially on Jensen's inequality. In the off-critical case, Jensen's inequality produces new m -dependent constants in both the rate c and the factor C , since $(\rho - m)_\pm$ are not probability distributions. We expect, however, that a more careful analysis would still provide uniform decay rates, in analogy to the formally equivalent statement for the MKR mixing measures $D(\rho)$. The constant C , however, will certainly show an m -dependence as a consequence of the sharp estimate [4, Lemma 1.2].*
3. *The result in Theorem 2 can be extended to smooth functions $\rho \in \mathbb{R}$ in several ways. One possibility is by replacing the BV norm by the Ginzburg–Landau energy*

$$E(\rho) = \int \frac{1}{2} |\nabla \rho|^2 + \frac{1}{2} (1 - \rho^2)^2 dx,$$

and restricting to the low-energy regime $E(\rho) \leq 1$ (which includes perfect mixing $\rho \equiv 0$), see also [16, Prop. 4.4]. In a certain rigorous sense (Γ -convergence), the Ginzburg–Landau energy is a smooth-interface regularization of the BV norm [14].

In the remainder of this paper, we prove Theorems 1 and 2.

3 Proofs

3.1 Proof of Theorem 1

Though the proof of Theorem 1 is a small variant of the proof of Proposition 2.2 in [2], we will provide most of its details for the sake of completeness and for the reader's convenience.

Proof of Theorem 1. Since u is divergence-free and $\rho_\pm = \pm(\rho \pm 1)$, both ρ_+ and ρ_- satisfy a continuity equation

$$\partial_t \rho_\pm + \nabla \cdot (u \rho_\pm) = 0.$$

Arguing as in [2, Lemma 4.1], we can thus show that $t \mapsto d_{\ln_\varepsilon}(\rho_+(t), \rho_-(t))$ is absolutely continuous and

$$\frac{d}{dt}d_{\ln_\varepsilon}(\rho_+, \rho_-) \gtrsim - \iint \frac{|u(x) - u(y)|}{|x - y|} d\pi_*^\varepsilon(t)(x, y),$$

where $\ln_\varepsilon(z) = \ln(z + \varepsilon) - \ln \varepsilon$ for some small $\varepsilon > 0$ and $\pi_*^\varepsilon(t)$ denotes an optimal transport plan for the MKR distance $d_{\ln_\varepsilon}(\rho_+(t), \rho_-(t))$. Here we have introduced a regularization of the cost function in order to ensure finiteness for all $z \geq 0$. The idea how to estimate the above Lipschitz norm of u by the L^p norm of its gradient is inspired by the work of Crippa & DeLellis [5] and uses the notion of maximal functions. If $M := M_1^0$ denotes the maximal function operator, see Definition 3, it holds

$$\frac{|f(x) - f(y)|}{|x - y|} \lesssim (M|\nabla f|)(x) + (M|\nabla f|)(y), \quad (4)$$

and, for $p \in (1, \infty]$

$$\|M|f|\|_{L^p} \lesssim \|f\|_{L^p}. \quad (5)$$

The first estimate is proved in a slightly more general version in Lemma 1 of the appendix. The second estimate is the fundamental inequality for maximal functions, cf. [19, p. 5, Theorem 1]. Applying (4), we infer

$$\frac{d}{dt}d_{\ln_\varepsilon}(\rho_+, \rho_-) \gtrsim - \iint (M|\nabla u|)(x) + (M|\nabla u|)(y) d\pi_*^\varepsilon(t)(x, y),$$

which can be rewritten as

$$\frac{d}{dt}d_{\ln_\varepsilon}(\rho_+, \rho_-) \gtrsim - \int (M|\nabla u|)\rho_+ dx + \int (M|\nabla u|)\rho_- dy$$

by the definition of the transport plans. Now, using $\|\rho_\pm\|_{L^\infty} = 2$, Jensen's inequality and (5), we infer that

$$\frac{d}{dt}d_{\ln_\varepsilon}(\rho_+, \rho_-) \gtrsim -\|\nabla u\|_{L^p}.$$

Integrating in time now yields

$$d_{\ln_\varepsilon}(\rho_+(T), \rho_-(T)) \geq d_{\ln_\varepsilon}((\rho_0)_+, (\rho_0)_-) - c \int_0^T \|\nabla u\|_{L^p} dt,$$

for some constant $c > 0$ depending only on d and p . Since transport plans are probability measures, the above estimate implies

$$\iint \ln(|x - y| + \varepsilon) d\pi(T)(x, y) \geq \iint \ln(|x - y| + \varepsilon) d\pi_*^\varepsilon(0)(x, y) - c \int_0^T \|\nabla u\|_{L^p} dt,$$

for all $\pi(T) \in \Gamma(\rho_+(T), \rho_-(T))$. By the monotonicity of the logarithm, it is $\iint \ln(|x - y| + \varepsilon) d\pi_*^\varepsilon(0)(x, y) \geq d_{\ln}((\rho_0)_+, (\rho_0)_-)$, and thus, passing to the limit $\varepsilon \downarrow 0$ we have

$$\iint \ln|x - y| d\pi(T)(x, y) \geq d_{\ln}((\rho_0)_+, (\rho_0)_-) - c \int_0^T \|\nabla u\|_{L^p} dt,$$

for all $\pi(T) \in \Gamma(\rho_+(T), \rho_-(T))$. Minimizing over $\pi(T)$ and applying the exponential function yields the statement of the theorem. \square

3.2 Proof of Theorem 2

The statement in Theorem 2 is an immediate consequence of Theorem 1 and the following two lemmas:

Lemma 1. *For every periodic function $\rho : [0, 1]^d \rightarrow \mathbb{R}$ with $\int \rho dx = 0$, it holds that*

$$D(\rho) \leq \|\rho\|_{\dot{H}^{-1}}.$$

Lemma 2. *There exists a constant $c > 0$ depending only on the space dimension, such that for every $\rho : [0, 1]^d \rightarrow \{\pm 1\}$ with $\int \rho dx = 0$ it holds*

$$c \leq D(\rho)[\rho]_{BV}.$$

The first lemma shows that the MKR mixing measure is dominated by the \dot{H}^{-1} norm. The second lemma provides a lower bound on the MKR mixing measure in terms of the inverse BV norm and is derived as an interpolation inequality. In a certain sense, the estimate can be seen as a non-standard isoperimetric inequality.

We start with the relatively simple

Proof of Lemma 1. Let $\pi \in \Gamma(\rho_+, \rho_-)$ be arbitrary, so that

$$D(\rho) \leq \exp\left(\iint \ln|x-y| d\pi\right).$$

We use Jensen's inequality together with the fact that π is a probability measure to infer

$$D(\rho) \leq \iint |x-y| d\pi,$$

and thus, minimizing over all $\pi \in \Gamma(\rho_+, \rho_-)$ yields

$$D(\rho) \leq d_{id}(\rho_+, \rho_-).$$

By the Kantorovich–Rubinstein theorem (cf. [20, Theorem 1.14]), $d_{id}(\rho_+, \rho_-)$ is a dual norm, namely

$$d_{id}(\rho_+, \rho_-) = \sup\left\{\int \rho\zeta dx : \|\nabla\zeta\|_{L^\infty} \leq 1\right\}.$$

Because $\|\nabla\zeta\|_{L^2} \leq \|\nabla\zeta\|_{L^\infty}$, we thus deduce that

$$D(\rho) \leq \sup\left\{\int \rho\zeta dx : \|\nabla\zeta\|_{L^2} \leq 1\right\} = \|\rho\|_{\dot{H}^{-1}}.$$

□

The proof of the interpolation inequality of Lemma 2 is more involved. It comes as a nontrivial extension of the estimate [2, Prop. 2.3]. In the new proof below, we use *weighted local maximal functions*, that will be introduced and discussed in the appendix.

Proof of Lemma 2. Let $0 < r \leq 1$ and $R > 0$ be two arbitrary constants that will be fixed at the end of the proof. We denote by ρ_R the convolution of ρ with a radially symmetric standard mollifier that is supported on a ball of radius R . We then write

$$1 = \int \rho^2 dx = \int \rho(\rho - \rho_R) dx + \int \rho \rho_R dx. \quad (6)$$

The estimate of the first integral on the right is standard and can easily be derived via approximation by smooth functions, namely

$$\int \rho(\rho - \rho_R) dx \leq \|\rho\|_{L^\infty} R \int |\nabla \rho| = 2R[\rho]_{BV}. \quad (7)$$

For the second term, we choose an arbitrary transport plan $\pi \in \Gamma(\rho_+, \rho_-)$ and write

$$\begin{aligned} & \int \rho \rho_R dx \\ &= \frac{1}{2} \int \rho_R(\rho_+ - \rho_-) dx \\ &= \frac{1}{2} \iint (\rho_R(x) - \rho_R(y)) d\pi \\ &= \frac{1}{2} \iint_{|x-y| \geq r} (\rho_R(x) - \rho_R(y)) d\pi + \frac{1}{2} \iint_{|x-y| < r} (\rho_R(x) - \rho_R(y)) d\pi. \end{aligned} \quad (8)$$

Since π is a probability measure, the second term in (8) can be controlled as follows:

$$\iint_{|x-y| < r} (\rho_R(x) - \rho_R(y)) d\pi \leq \|\nabla \rho_R\|_{L^\infty} r \iint d\pi \lesssim \|\rho\|_{L^\infty} \frac{r}{R} = \frac{r}{R}. \quad (9)$$

It remains to estimate the first term in (8). We first notice that $1 \leq \ln|x-y| + \ln \frac{e}{r}$ for all $|x-y| \geq r$, and thus

$$\begin{aligned} & \iint_{|x-y| \geq r} (\rho_R(x) - \rho_R(y)) d\pi \\ & \leq 2\|\rho_R\|_{L^\infty} \iint_{|x-y| \geq r} d\pi \\ & \leq 4 \iint \left(\ln|x-y| + \ln \frac{e}{r} \right) d\pi + 4 \iint_{|x-y| < r} \ln \frac{r}{e|x-y|} d\pi \\ & = 4 \iint \ln|x-y| d\pi + 4 \ln \frac{e}{r} + 4 \iint_{|x-y| < r} \ln \frac{r}{e|x-y|} d\pi. \end{aligned} \quad (10)$$

We now have

$$\begin{aligned} & \iint_{|x-y| < r} \ln \frac{r}{e|x-y|} d\pi \\ &= \iint_{|x-y| < r} \ln \frac{e^2 r}{|x-y|} d\pi - 3 \iint_{|x-y| < r} d\pi \\ & \leq \frac{1}{2} \iint_{|x-y| < r} \sqrt{|x-y|} \left(\ln \frac{e^2 r}{|x-y|} \right) \frac{|\rho(x) - \rho(y)|}{\sqrt{|x-y|}} d\pi, \end{aligned}$$

because $\rho(x) = 1$ and $\rho(y) = -1$ for all (x, y) in the support of π . Observe that $s \mapsto \sqrt{s} \ln \frac{e^2 r}{s}$ is an increasing function on $(0, r]$, and thus

$$\iint_{|x-y|<r} \ln \frac{r}{e|x-y|} d\pi \leq \sqrt{r} \iint_{|x-y|<r} \frac{|\rho(x) - \rho(y)|}{\sqrt{|x-y|}} d\pi.$$

The integral on the right can be estimated with the help of weighted local maximal functions, that will be introduced and discussed in the appendix. For this purpose, let $\{\rho_\varepsilon\}_{\varepsilon \downarrow 0}$ be a smooth approximation of ρ . Applying Lemma 3 below with $f = \rho_\varepsilon$ and $\theta = \frac{1}{2}$ yields

$$\iint_{|x-y|<r} \frac{|\rho_\varepsilon(x) - \rho_\varepsilon(y)|}{\sqrt{|x-y|}} d\pi \lesssim \iint_{|x-y|<r} ((M_r^{1/2} |\nabla \rho_\varepsilon|)(x) + (M_r^{1/2} |\nabla \rho_\varepsilon|)(y)) d\pi,$$

and thus, by the definition of π and since $\|\rho_\pm\|_{L^\infty} = 2$

$$\iint_{|x-y|<r} \frac{|\rho_\varepsilon(x) - \rho_\varepsilon(y)|}{\sqrt{|x-y|}} d\pi \lesssim \int |M_r^{1/2} |\nabla \rho_\varepsilon|| dx.$$

We now invoke Lemma 4 with $f = |\nabla \rho_\varepsilon|$ and $\tau = \frac{1}{2}$, to the effect of

$$\iint_{|x-y|<r} \frac{|\rho_\varepsilon(x) - \rho_\varepsilon(y)|}{\sqrt{|x-y|}} d\pi \lesssim \sqrt{r} \int |\nabla \rho_\varepsilon| dx.$$

Passing to the limit $\varepsilon \downarrow 0$ via Fatou's lemma, the above calculation shows

$$\iint_{|x-y|<r} \ln \frac{r}{e|x-y|} d\pi \lesssim r[\rho]_{BV}. \quad (11)$$

Finally, combining (6)–(11) yields

$$1 \leq 2R[\rho]_{BV} + C \frac{r}{R} + 2 \iint \ln |x-y| d\pi + 2 \ln \frac{e}{r} + Cr[\rho]_{BV},$$

for some $C > 0$. We first choose $R = Cr$ to the effect of

$$0 \leq \iint \ln |x-y| d\pi + \ln \frac{e}{r} + \frac{3}{2} Cr[\rho]_{BV}.$$

Now, optimizing in r yields $r \sim \frac{1}{[\rho]_{BV}}$ (observe that $[\rho]_{BV} \gtrsim 1$ by the virtue of the isoperimetric inequality) and obtain, upon applying the exponential function to both sides of the inequality

$$c \leq \exp \left(\iint \ln |x-y| d\pi \right) [\rho]_{BV},$$

for some new constant $c > 0$. The statement of the lemma follows when optimizing over $\pi \in \Gamma(\rho_+, \rho_-)$. \square

Appendix: Weighted local minimal functions

The proof of Lemma 2 above uses the notion of weighted local maximal functions that will be introduced and discussed in the following.

Definition 3 (Weighted local maximal functions). *Let $f : [0, 1]^d \rightarrow [0, \infty)$ denote a smooth periodic function, and let $0 < r \leq 1$ and $0 \leq \tau \leq 1$ be two constants. We define the weighted local maximal function of f as*

$$(M_r^\tau f)(x) := \sup_{0 < q < r} \frac{1}{|B_q(0)|} \int_{B_q(0)} |z|^\tau f(x+z) dx. \quad (12)$$

Observe that we recover classical local maximal functions in the case $\tau = 0$, cf. [19, Ch. 1]. The following estimate is well-known to hold in the case $\theta = 1$ (or $\tau = 0$).

Lemma 3. *Let $f : [0, 1]^d \rightarrow \mathbb{R}$ denote a smooth periodic function, and let $0 < r \leq 1$ and $0 < \theta \leq 1$ be two constants. Then there exists a constant $C > 0$ depending only on d and θ such that*

$$\frac{|f(x) - f(y)|}{|x - y|^\theta} \leq C ((M_r^{1-\theta} |\nabla f|)(x) + (M_r^{1-\theta} |\nabla f|)(y)), \quad (13)$$

for all $x, y \in [0, 1]^d$ such that $|x - y| \leq r$.

Proof. Let $Q := |x - y|$. We start with a classical argument that replaces a pointwise by an averaged quantity: We first show that

$$\frac{|f(x) - f(y)|}{|x - y|^\theta} \lesssim \frac{1}{Q^d} \int_{B_Q(x)} \frac{|f(x) - f(z)|}{|x - z|^\theta} dz + \frac{1}{Q^d} \int_{B_Q(y)} \frac{|f(y) - f(z)|}{|y - z|^\theta} dz.$$

Indeed, via the triangle inequality and averaging over $B_Q(x) \cap B_Q(y)$, we see that

$$\begin{aligned} & |f(x) - f(y)| \\ & \lesssim \frac{1}{Q^d} \int_{B_Q(x) \cap B_Q(y)} |f(x) - f(z)| dz + \frac{1}{Q^d} \int_{B_Q(x) \cap B_Q(y)} |f(y) - f(z)| dz \\ & \leq \frac{1}{Q^d} \int_{B_Q(x)} |f(x) - f(z)| dz + \frac{1}{Q^d} \int_{B_Q(y)} |f(y) - f(z)| dz, \end{aligned}$$

and the statement follows from $|x - z| \leq Q = |x - y|$ for all $z \in B_Q(x)$. For the statement in (12) it thus suffices to show that

$$\frac{1}{Q^d} \int_{B_Q(x)} \frac{|f(x) - f(z)|}{|x - z|^\theta} dz \lesssim (M_r^{1-\theta} |\nabla f|)(x), \quad (14)$$

for all $x \in [0, 1]^d$. Without loss of generality, we may assume that $x = 0$. Then, using the Fundamental Theorem of Calculus, we estimate

$$\frac{1}{Q^d} \int_{B_Q(0)} \frac{|f(0) - f(z)|}{|z|^\theta} dz \leq \frac{1}{Q^d} \int_{B_Q(0)} |z|^{1-\theta} \int_0^1 |\nabla f(sz)| ds dz.$$

Furthermore, via Fubini and a change of variables, this becomes

$$\begin{aligned} \frac{1}{Q^d} \int_{B_Q(0)} \frac{|f(0) - f(z)|}{|z|^\theta} dz &\leq \int_0^1 \frac{1}{s^{1-\theta}} \frac{1}{(sQ)^d} \int_{B_{sQ}(0)} |\hat{z}|^{1-\theta} |\nabla f(\hat{z})| d\hat{z} ds \\ &\leq \left(\int_0^1 \frac{1}{s^{1-\theta}} ds \right) (M_r^{1-\theta} |\nabla f|)(0), \end{aligned}$$

where in the last inequality we have used $sQ \leq sr \leq r$ and the definition of weighted local maximal functions. In order to deduce (14), it remains to observe that the prefactor is bounded for $\theta > 0$. \square

We finally derive the fundamental inequality for weighted local maximal functions. Observe that this estimate fails in the case $\tau = 0$, that is, for classical maximal functions.

Lemma 4. *Let $f : [0, 1]^d \rightarrow [0, \infty)$ denote a smooth periodic function, and let $0 < r \leq 1$ and $0 < \tau \leq 1$ be two constants. Then there exists a constant $C > 0$ depending only on d and τ such that*

$$\int |M_r^\tau f| dx \leq Cr^\tau \int |f| dx. \quad (15)$$

Proof. For every $0 < q \leq r$, we define

$$F(x, q) := \frac{1}{q^d} \int_{B_q(0)} |z|^\tau f(x+z) dz.$$

Because $\tau > 0$, we have that $\lim_{|q| \downarrow 0} F(\cdot, q) = 0$, and thus the Fundamental Theorem of Calculus yields

$$\begin{aligned} F(x, q) &= \int_0^q \frac{\partial}{\partial Q} F(x, Q) dQ \\ &= \int_0^q \left(-\frac{d}{Q^{d+1}} \int_{B_Q(0)} |z|^\tau f(x+z) dz + \frac{1}{Q^d} \int_{\partial B_Q(0)} |z|^\tau f(x+z) d\mathcal{H}^{d-1}(z) \right) dQ \\ &\leq \int_0^r \frac{1}{Q^d} \int_{\partial B_Q(0)} |z|^\tau f(x+z) d\mathcal{H}^{d-1}(z) dQ, \end{aligned}$$

where in the last inequality we have used that $q \leq r$ and $f \geq 0$. Via a change of variables $z = Q\hat{z}$, we obtain the estimate

$$F(x, q) \lesssim \int_0^r Q^{\tau-1} \int_{\partial B_1(0)} f(x+Q\hat{z}) d\mathcal{H}^{d-1}(\hat{z}) dQ.$$

We take the supremum over all $0 < q \leq r$ on the left hand side and integrate over $[0, 1)^d$. Using Fubini and the periodicity of f , we obtain

$$\begin{aligned} \int |M_r^\tau f| dx &\lesssim \int \int_0^r Q^{\tau-1} \int_{\partial B_1(0)} f(x + Q\hat{z}) d\mathcal{H}^{d-1}(\hat{z}) dQ dx \\ &= \int_0^r Q^{\tau-1} dQ \int_{\partial B_1(0)} d\mathcal{H}^{d-1}(\hat{z}) \int |f| dx \\ &\lesssim r^\tau \int |f| dx. \end{aligned}$$

Notice that we have used $\tau > 0$ in the last estimate. This proves Lemma 4 □

Acknowledgment

The author thanks Charlie Doering and Evelyn Lunasin for bringing this topic to his attention, for enlightening discussions, and for the kind hospitality at the University of Michigan, Ann Arbor. He also acknowledges motivating discussions with Felix Otto.

References

- [1] ALBERTI, G., CHOKSI, R., AND OTTO, F. Uniform energy distribution for an isoperimetric problem with long-range interactions. *J. Amer. Math. Soc.* *22* (2009), 569–605.
- [2] BRENIER, Y., OTTO, F., AND SEIS, C. Upper bounds on coarsening rates in demixing binary viscous liquids. *SIAM J. Math. Anal.* *43*, 1 (2011), 114–134.
- [3] BRESSAN, A. A lemma and a conjecture on the cost of rearrangements. *Rend. Sem. Mat. Univ. Padova* *110* (2003), 97–102.
- [4] CONTI, S., NIETHAMMER, B., AND OTTO, F. Coarsening rates in off-critical mixtures. *SIAM J. Math. Anal.* *37*, 6 (2006), 1732–1741 (electronic).
- [5] CRIPPA, G., AND DE LELLIS, C. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.* *616* (2008), 15–46.
- [6] DOERING, C. R., AND THIFFEAULT, J.-L. Multiscale mixing efficiencies for steady sources. *Phys. Rev. E* *74* (Aug 2006), 025301.
- [7] GUBANOV, O., AND CORTELZZI, L. Towards the design of an optimal mixer. *Journal of Fluid Mechanics* *651* (5 2010), 27–53.
- [8] KOHN, R. V., AND OTTO, F. Upper bounds on coarsening rates. *Comm. Math. Phys.* *229*, 3 (2002), 375–395.

- [9] LIN, Z., THIFFEAULT, J.-L., AND DOERING, C. R. Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.* 675 (2011), 465–476.
- [10] LIU, W. Mixing enhancement by optimal flow advection. *SIAM J. Control Optim.* 47, 2 (2008), 624–638.
- [11] LUNASIN, E., LIN, Z., NOVIKOV, A., MAZZUCATO, A., AND DOERING, C. R. Optimal mixing and optimal stirring for fixed energy, fixed power, or fixed palenstrophy flows. *J. Math. Phys.* 53, 11 (2012), 115611, 15.
- [12] MATHEW, G., MEZIĆ, I., GRIVOPOULOS, S., VAIDYA, U., AND PETZOLD, L. Optimal control of mixing in Stokes fluid flows. *J. Fluid Mech.* 580 (2007), 261–281.
- [13] MATHEW, G., MEZIĆ, I., AND PETZOLD, L. A multiscale measure for mixing. *Phys. D* 211, 1-2 (2005), 23–46.
- [14] MODICA, L., AND MORTOLA, S. Un esempio di Γ^- -convergenza. *Boll. Un. Mat. Ital. B (5)* 14, 1 (1977), 285–299.
- [15] OTTO, F., RUMP, T., AND SLEPČEV, D. Coarsening rates for a droplet model: rigorous upper bounds. *SIAM J. Math. Anal.* 38, 2 (2006), 503–529 (electronic).
- [16] OTTO, F., SEIS, C., AND SLEPČEV, D. Crossover of the coarsening rates in demixing of binary viscous liquids. *Commun. Math. Sci.* 11, 2 (2013), 441–464.
- [17] SHAW, T. A., THIFFEAULT, J.-L., AND DOERING, C. R. Stirring up trouble: Multi-scale mixing measures for steady scalar sources. *Physica D: Nonlinear Phenomena* 231, 2 (2007), 143 – 164.
- [18] SLEPČEV, D. Coarsening in nonlocal interfacial systems. *SIAM J. Math. Anal.* 40, 3 (2008), 1029–1048.
- [19] STEIN, E. M. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.
- [20] VILLANI, C. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [21] VILLANI, C. *Optimal transport*, vol. 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.