Vanishing geodesic distance for right-invariant Sobolev metrics on diffeomorphism groups

Robert L. Jerrard* and Cy Maor*

Abstract

We study the geodesic distance induced by right-invariant metrics on the group $\text{Diff}_c(M)$ of compactly supported diffeomorphisms, for various Sobolev norms $W^{s,p}$. Our main result is that the geodesic distance vanishes identically on every connected component whenever $s < \min\{n/p, 1\}$, where $n$ is the dimension of $M$. We also show that previous results imply that whenever $s > n/p$ or $s \geq 1$, the geodesic distance is always positive. In particular, when $n \geq 2$, the geodesic distance vanishes if and only if $s < 1$ in the Riemannian case $p = 2$, contrary to a conjecture made in [BBHM13].

Contents

1 Introduction

2 Preliminaries and main result

3 Two-dimensional construction

3.1 Step I: Splitting into strips

3.2 Step II: Squeezing the strips

3.3 Step III: Flowing along the squeezed strips

3.4 Step IV: Error correction – affine homotopy

4 Higher-dimensional construction

4.1 Step I: Splitting into strips

4.2 Step II: Squeezing the strips

4.3 Step III: Affine homotopy

5 The construction for $W^{s,p}(\mathbb{R}^n)$, for $n \geq 2$. 

*Department of Mathematics, University of Toronto.
1 Introduction

In this paper we mostly resolve a question about the geometry of the group $\text{Diff}_c(M)$ of compactly supported diffeomorphisms of a Riemannian manifold $M$, endowed with a right-invariant Sobolev metric; see Section 2 below for the precise definition, as well as assumptions on $M$. Sobolev metrics on $\text{Diff}_c(M)$ arise in a variety of contexts. In particular, such a metric turns $\text{Diff}_c(M)$ into an infinite-dimensional Riemannian manifold, and a number of partial differential equations relevant to fluid dynamics can be formulated as geodesic flow in manifolds of this sort. Sobolev metrics on $\text{Diff}_c(M)$ are also relevant to the study of what are known as shape spaces, a concept with connections to areas such as computer vision and computational anatomy. We refer to [BBM14] for a discussion of these and other sources of motivation.

A metric on $\text{Diff}_c(M)$ gives rise to a notion of the length of a path, and the induced geodesic distance between a pair of elements is obtained by taking the infimum of the lengths of all paths connecting the two diffeomorphisms. If the metric is induced by the $H^s$ Sobolev inner product for $s$ small enough, the geodesic distance may vanish in the strong sense that any two diffeomorphisms that can be connected by a path can in fact be connected by a path of arbitrarily small length. For large enough $s$, by contrast, the geodesic distance between any two distinct diffeomorphisms is positive. Our aim is to identify the precise threshold that separates these two cases.

This question grows out of work of [MM05], who proved (among other results) that the $H^s$ geodesic distance vanishes when $s = 0$ and is positive when $s = 1$. These results were extended to certain $s \in (0, 1)$ by [BBHM13, BBM13], who proved that for $M$ of bounded geometry, the $H^s$ geodesic distance vanishes if $s < 1/2$. They also proved that for 1-dimensional manifolds, the geodesic distance is positive when $s > 1/2$, and for $M = S^1$, it vanishes in the borderline case $s = 1/2$. Motivated by these facts, they conjectured that for arbitrary manifolds, the induced $H^s$ geodesic distance should vanish if and only if $s \leq 1/2$.

It turns out to be illuminating to embed this conjecture in a larger family of questions, about the vanishing of the right-invariant geodesic distance induced by fractional Sobolev norms $W^{s,p}$, for $1 \leq p \leq \infty$, see again Section 2 for details. The arguments used by [MM05, Theorem 5.7], [BBHM13, Theorem 4.1] then imply the following:

**Theorem 1.1** ([MM05, BBHM13]) The induced $W^{s,p}$-distance is positive whenever $sp > n$ or $s \geq 1$.

Our main result shows that these results are essentially sharp:

**Theorem 1.2** The induced $W^{s,p}$-distance is vanishes whenever $sp < n$ and $s < 1$.

These result are stated in a more detailed way in Theorem 2.4. In particular, contrary to the conjecture of [BBHM13], we have the following corollary:

**Corollary 1.3** If $M$ is a manifold of dimension at least 2, then the $H^s$ geodesic distance vanishes if and only if $s < 1$.

We conclude this informal introduction by describing some ingredients in our analysis. First, we remark that the positivity proof of [MM05, Theorem 5.7] can be understood to show that for any $s \geq 0$, paths in $\text{Diff}_c(M)$ of short length can only be obtained by compressing the support of the diffeomorphisms into very small sets, and that this compression can
always be detected by $W^{s,p}$-norms when $s \geq 1$. The positivity proof of [BBHM13 Theorem 4.1] relies on the observation that any motion, no matter how small its support, can always detected by any $W^{s,p}$-norm that embeds into $L^\infty$. This property holds whenever $sp > n$.

If $s < 1$, it turns out that one can compress parts of the manifold into arbitrarily small regions, for arbitrarily small cost; and if $sp < n$ one can transport small regions of the manifold for a long distance with small cost. Therefore, if $s < \min\{n/p, 1\}$, one might expect the geodesic distance to vanish. Our proof that this is indeed the case has two main points. The first is to devise a strategy for alternating compression and transport of small sets in order to flow the identity mapping, say, onto a fixed target di ff erent from the desired target; however in order for this flow to be in the right Sobolev space we need to regularize it. This regularization, and the error controlling that follows it, form the majority of the technical part of this paper.

Our heuristic arguments, described above, for vanishing geodesic distance apply also in the endpoint case $s = \frac{n}{p} < 1$, since $W^{n/p,p}$ also fails to embed into $L^\infty$ in this case. As mentioned above, it is known that the $W^{1/2,2}$-induced geodesic distance vanishes on $\text{Diff}_c(S^1)$, and although we do not present the details, the proof of [BBHM13] can be readily extended to $W^{1/p,p}$ for all $1 < p < \infty$. In general, however, although it is natural to conjecture that the $W^{n/p,p}$-induced geodesic distance vanishes on $n$ dimensional manifolds when $p > n$, the critical scaling makes constructions delicate, and this question remains open except when $\mathcal{M} = S^1$.

2 Preliminaries and main result

Let $(\mathcal{M}, g)$ be a Riemannian manifold of bounded geometry, that is $(\mathcal{M}, g)$ has a positive injectivity radius and all the covariant derivatives of the curvature are bounded: $\|\nabla^i R\|_g < C_i$ for $i \geq 0$. We denote by $\Gamma_c(TM)$ the Lie-algebra of compactly supported vector field on $\mathcal{M}$, and by $\text{Diff}_c(\mathcal{M})$ the group of compactly supported diffeomorphisms of $\mathcal{M}$, that is the diffeomorphisms $\phi$ for which the closure of $\{\phi(x) \neq x\}$ is compact.

A smooth path $\{\phi_t\}_{t \in [0,1]}$ in $\text{Diff}_c(\mathcal{M})$ can be described in terms of the velocity vector fields $u(t, \cdot)$ such that $\partial_t \phi_t = u(t, \phi_t)$ for $0 \leq t \leq 1$. Given $\{\phi_t\}$, we find $u$ by setting $u(t, \cdot) := \partial_t \phi_t \circ \phi_t^{-1}$, and conversely $\{\phi_t\}_{t \in [0,1]}$ may be recovered from $u$ and $\phi_0$ by standard ODE theory. Given a norm $\| \cdot \|_A$ on $\Gamma_c(TM)$ we can then define the geodesic distance between $\phi_0, \phi_1 \in \text{Diff}_c(\mathcal{M})$ by

$$\text{dist}_A(\phi_0, \phi_1) := \inf \left\{ \int_0^1 \|u(t)\|_A \, dt : \partial_t \phi_t = u(t, \phi_t) \text{ for } 0 \leq t \leq 1 \right\}.$$  

Note that $\text{dist}_A$ forms a semi-metric on $\text{Diff}_c(\mathcal{M})$, that is it satisfies the triangle inequality but may fail to be positive.

This is the geodesic distance of the right-invariant Finsler metric on $\text{Diff}_c(\mathcal{M})$ induced by $\| \cdot \|_A$, which is defined as

$$\|X\|_{A, \phi} := \|X \circ \phi^{-1}\|_A$$

3
for every \( \phi \in \text{Diff}_c(M) \) and \( X \in T_0 \text{Diff}_c(M) \). If \( \| \cdot \|_A \) comes from an inner-product, it defines a Riemannian metric on \( \text{Diff}_c(M) \) in a similar manner. See [BBHM13] for more details. The right-invariance of \( \text{dist}_A \) is summarized in the following lemma:

**Lemma 2.1 (Right-invariance)** For \( \psi, \phi_0, \phi_1 \in \text{Diff}_c(M) \), we have

\[
\text{dist}_A(\phi_0 \circ \psi, \phi_1 \circ \psi) = \text{dist}_A(\phi_0, \phi_1).
\]

In particular,

\[
\text{dist}_A(\text{Id}, \psi) = \text{dist}_A(\text{Id}, \psi^{-1}),
\]

and

\[
\text{dist}_A(\text{Id}, \phi_1 \circ \phi_0) \leq \text{dist}_A(\text{Id}, \phi_1) + \text{dist}_A(\text{Id}, \phi_0).
\]

**Proof:** Let \( t \mapsto \phi_t \in \text{Diff}_c(M) \) be a curve from \( \phi_0 \) to \( \phi_1 \). Denote \( u_t = \partial_t \phi_t \circ \phi_t^{-1} \). Define \( \Phi_t = \phi_t \circ \psi \). This is a curve from \( \phi_0 \circ \psi \) to \( \phi_1 \circ \psi \). We then have

\[
\partial_t \Phi_t = \partial_t \phi_t \circ \psi = \partial_t \phi_t \circ \phi_t^{-1} \circ \Phi_t = u_t \circ \Phi_t,
\]

from which the first claim follows immediately. The second and third claims follow from the first, since

\[
\text{dist}_A(\text{Id}, \psi^{-1}) = \text{dist}_A(\psi \circ \psi^{-1}, \psi^{-1}) = \text{dist}_A(\psi, \text{Id}),
\]

and

\[
\text{dist}_A(\text{Id}, \phi_1 \circ \phi_0) \leq \text{dist}_A(\text{Id}, \phi_0) + \text{dist}_A(\phi_0, \phi_1) = \text{dist}_A(\text{Id}, \phi_0) + \text{dist}_A(\text{Id}, \phi_1).
\]

We are interested in fractional Sobolev \( W^{s,p} \)-norms, and in particular in \( H^s := W^{s,2} \), for \( s \in (0,1) \). We adopt the following as our basic definition, from among a number of equivalent formulations.

**Definition 2.2** For \( 0 < s < 1 \) and \( 1 \leq p < \infty \), the \( W^{s,p} \)-norm of a function \( f \in L^p(\mathbb{R}^n) \) is given by

\[
\|f\|_{s,p}^p = \|f\|_{L^p}^p + \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|f(x) - f(y)|^p}{|x - y|^{n+sp}} \, dx \, dy.
\]

Given a Riemannian manifold \( (M, g) \) of bounded geometry, this norm can be extended to \( \Gamma_c(TM) \) using trivialization by normal coordinate patches on \( M \) (see [BBM13] Section 2.2) for details). We will denote the induced geodesic distance on \( \text{Diff}_c(M) \) by \( \text{dist}_{s,p} \). When \( p = 2 \), we will denote \( \text{dist}_{s,2} \) by \( \text{dist}_s \) for simplicity. Different choices of charts result in equivalent metrics, and therefore the question of vanishing geodesic distance is independent of these choices.

Instead of using Definition 2.2 directly, we will bound the \( W^{s,p} \)-norm using an interpolation inequality:

**Proposition 2.3 (fractional Gagliardo-Nirenberg interpolation inequality)** Assume that \( 1 < p < \infty \). For every \( f \in W^{1,p}(\mathbb{R}^n) \) and \( s \in (0,1) \),

\[
\|f\|_{s,p} \leq C_{s,p} \|f\|_{L^p}^{1-s} \|f\|_{1,p}^s,
\]

where

\[
\|f\|_{1,p}^p := \|f\|_{L^p}^p + \|df\|_{L^p}^p.
\]
For a proof, see for example [BM01, Corollary 3.2]. In fact this is the only property of the $W^{p,p}$-norm that we will use. We remark that when $p = 2$, the above inequality (with $C = 1$) follows immediately from Hölder’s inequality, if one uses the equivalent norm $\|f\|_{s,2}^2 = \int_{\mathbb{R}} (1 + |\xi|^2)^{s/2} |\hat{f}(\xi)|^2 d\xi$, where $\hat{f}$ denotes the Fourier transform.

The main result of this paper is the following.

**Theorem 2.4** Let $(M, g)$ be an $n$-dimensional Riemannian manifold of bounded geometry.

1. If $p \geq 1$ and $s < \min\{1, n/p\}$, then $\text{dist}_{s,p}(\phi_0, \phi_1) = 0$ whenever $\phi_0, \phi_1$ belong to the same path-connected component of $\text{Diff}_c(M)$.

2. If $s \geq 1$ or $sp > n$ then $\text{dist}_{s,p}(\phi_0, \phi_1) > 0$ for any two distinct $\phi_0, \phi_1 \in \text{Diff}_c(M)$.

The second assertion is a direct consequence of known arguments in the case $p = 2$. So is the first one for the case $n = 1$. The new point is the vanishing of geodesic distance for all $s < \min\{1, n/p\}$ whenever $n \geq 2$.

In the remainder of this section we quickly verify that known results about the case $p = 2$ extend to the more general setting we consider here, and we present the reduction, also well-known in the $H^s$ case, that will allow us to complete the proof of the theorem by showing that $\text{dist}_{s,p}(\text{Id}, \Phi) = 0$ for a single compactly supported diffeomorphism on $\mathbb{R}^n$.

**Positive geodesic distance.** First, assume that $\phi_0, \phi_1$ are two distinct elements of $\text{Diff}_c(M)$, and let $u$ be any time-dependent vector field generating a path $\phi : [0, 1] \to \text{Diff}_c(M)$ connecting $\phi_0$ to $\phi_1$, via the ODE $\partial_t \phi_t = u(t, \phi_t), 0 < t < 1$. The proof of [MM05, Theorem 5.7] uses a clever integration by parts to show that for any $\rho, \zeta \in C^1_c(M)$,

$$\left| \int_M \rho(\zeta \circ \psi_1 - \zeta) \text{vol}(g) \right| = \left| \int_0^1 \int_M (\zeta \circ \psi_t) \text{div}(\rho u) \text{vol}(g) \right|, \quad \psi_t := \phi_0 \circ \phi_t^{-1}.$$

By a suitable choice of $\rho, \zeta$, one finds that $0 < c \leq C \int_0^1 \|u(t)\|_{1,p} dt$ for $p \geq 1$, where the constants depend on $\phi_1, \phi_2, \rho, \zeta, p$. This shows the positivity of the geodesic distance in $W^{1,p}$ for any $p \geq 1$, and hence (since these spaces embed into $W^{1,p}$) in $W^{s,p}$ for $s \geq 1$.

On the other hand, if $s > n/p$, then $W^{s,p}$ embeds into some $C^{0,\alpha}$ (see for example [NPV12, Theorem 8.2]) and hence into $L^\infty$. Thus $\|\partial_t \phi_t\|_{L^\infty} = \|u(t)\|_{L^\infty} \leq C\|u(t)\|_{s,p}$, and as noted in [BBHM13, Theorem 4.1], the positivity of $\text{dist}_{s,p}$ follows directly:

$$|\phi_1(x) - \phi_0(x)| = \left| \int_0^1 \partial_t \phi_t(x) dt \right| \leq C \int_0^1 \|u(t)\|_{s,p} dt \quad \text{for every } x \in M.$$

For $sp < n = 1$, the proof of vanishing geodesic distance in [BBHM13] in the case $p = 2$ relies on an explicit construction (incorporated into (3.10) below) of a transportation scheme of the identity to a single diffeomorphism, that has arbitrarily small cost; this arbitrarily small cost follows from the fact that the $W^{s,p}$-norm of the characteristic function of an interval tends to zero with the length of the interval. For general $sp < n = 1$, this is well-known and can easily be verified from Definition 2.2. Once this is noted, the proof goes through with no change.
Reduction to a single diffeomorphism. The following proposition states an important property of \((\text{Diff}_c(M), \text{dist}_{s,p})\) – it is either a metric space, or it collapses completely, that is, the geodesic distance in any connected component of \(\text{Diff}_c(M)\) vanishes. In other words, if \((\text{Diff}_c(M), \text{dist}_{s,p})\) is not a metric space, then any two diffeomorphisms in the same connected component can be connected by a path of arbitrary short \(W^{s,p}\)-length.

**Proposition 2.5** Denote by \(\text{Diff}_0(M)\) the connected component of the identity (all diffeomorphisms in \(\text{Diff}_c(M)\) for which there exists a curve between them and Id).

1. \(\text{Diff}_0(M)\) is a simple group.

2. \(\left\{ \phi : \text{dist}_{s,p}(\text{Id}, \phi) = 0 \right\}\) is a normal subgroup of \(\text{Diff}_0(M)\). Therefore, it is either \{Id\} or the whole \(\text{Diff}_0(M)\).

This is proved in [BBHM13, p. 15] (see also [BBM14, Lemma 7.10]) when \(p = 2\), and the proof goes through with no essentially no change in our setting. We recall the idea. The first conclusion is classical (and is independent of the norm). To establish the second, we consider \(\phi, \psi \in \text{Diff}_c(M)\) such that \(\text{dist}_{s,p}(\text{Id}, \phi) = 0\), and we must show that \(\text{dist}_{s,p}(\text{Id}, \Phi) = 0\) for \(\Phi := \psi^{-1} \circ \phi \circ \psi\). To do this, note that if \(\phi_t, 0 \leq t \leq 1\) is a path connecting \(\text{Id}\) to \(\phi\), then \(\Phi_t := \psi^{-1} \circ \phi_t \circ \psi\) connects \(\text{Id}\) to \(\Phi\). The conclusion thus follows by verifying that

\[
\int_0^1 \| \partial_t \Phi_t \circ \Phi^{-1} \|_{s,p} dt \leq C \int_0^1 \| \partial_t \phi_t \circ \phi^{-1} \|_{s,p} dt,
\]

where \(C\) may depend on \(\psi, (M, g)\), \(s, p\) but not \(\phi\). In fact a pointwise inequality of the integrands holds for every \(t\). This follows after a computation from the fact that for \(h \in C^\infty(M)\) and \(\psi \in \text{Diff}_c(M)\), the operations of pointwise multiplication \(u \mapsto h \cdot u\) and composition \(u \mapsto u \circ \psi\) are bounded linear operators on \(W^{s,p}(M)\), see Theorems 4.2.2 and 4.3.2 in [Tri92].

The strategy for proving vanishing geodesic distance. The proof of part 1 of Theorem 2.4 for \(n \geq 2\) goes as follows:

1. For \(sp < n\) and \(n \geq 2\), we will show that there exists at least one nontrivial \(\Phi \in \text{Diff}_c(\mathbb{R}^n)\) such that \(\text{dist}_{s,p}(\text{Id}, \Phi) = 0\).

2. For general \((M, g)\) of bounded geometry, we can push-forward this example in \(\mathbb{R}^n\) to obtain a diffeomorphism \(\widetilde{\Phi}\), supported in a single coordinate chart used in the definition of induced \(W^{s,p}\) geodesic distance. Then the definitions imply that \(\text{dist}_{s,p}(\text{Id}, \widetilde{\Phi}) = 0\). (see [BBM13] for a similar argument).

3. Part 1 of Theorem 2.4 then follows from Proposition 2.5.

In the rest of the paper we treat the first point. For simplicity, we first consider the special case \(p = 2, M = \mathbb{R}^2\), and we show that \(\text{dist}_{c}(\text{Id}, \Phi) := \text{dist}_{s,2}(\text{Id}, \Phi) = 0\) for a particular \(\Phi \in \text{Diff}_c(\mathbb{R}^2)\). This construction, carried out in Section 3, contains all the ingredients of more general cases. In Section 4 we present a much simpler construction that works when \(p = 2, s < 1\) and \(n \geq 3\), and in Section 5 we show how to modify these arguments to complete the proof of the theorem in the general case.
3 Two-dimensional construction

In this section we prove the following:

**Theorem 3.1** Let \( \zeta \in C^\infty_c((0,1)^2) \) satisfying \( \zeta \geq 0, \partial_1 \zeta > -1 \). Denote \( \phi(x, y) = x + \zeta(x, y) \), and define \( \Phi \in \text{Diff}_c(\mathbb{R}^2) \) by \( \Phi(x, y) = (\phi(x, y), y) \). Then \( \text{dist}_s(\Phi, \text{Id}) = 0 \) for every \( s \in [0, 1) \).

We start with a general outline of the proof. Fix \( k \in \mathbb{N} \). In Section 3.1 we decompose \( \Phi \) as follows:

\[
\Phi = \Phi_2 \circ \Phi_1, \quad \Phi_i = (\phi_i(x, y), y) = (x + \zeta_i(x, y), y) \in \text{Diff}_c(\mathbb{R}^2),
\]

where \( \zeta_i \) is supported on the union of \( \approx k \) strips \((0,1) \times I_j, \|I_j\| \approx k^{-1}\). In Sections 3.2-3.4, we show that \( \text{dist}_s(\Phi_1, \text{Id}) = o(1) \), when \( k \to \infty \); the proof for \( \Phi_2 \) is analogous, and since \( k \) is arbitrary, the conclusion \( \text{dist}_s(\Phi, \text{Id}) = 0 \) follows by Lemma 2.1.

In order to prove \( \text{dist}_s(\Phi_1, \text{Id}) = o(1) \), we decompose \( \Phi_1 \) as follows:

\[
\Phi_1 = \Gamma^{-1} \circ \Psi^{-1} \circ \Theta \circ \Psi, \quad \Gamma, \Theta, \Psi \in \text{Diff}_c(\mathbb{R}^2),
\]

where

1. \( \Psi(x, y) = (x, \psi(x, y)) \) squeezes the intervals \( I_j \) into intervals of length \( \approx \lambda \) for \( \lambda \) of the form \( \lambda = e^{-\alpha k^{-1}} \), where \( \alpha = \alpha(k) \) is a (moderately large) parameter, to be determined. In Section 3.2 we define \( \Psi \) and show that \( \text{dist}_s(\Psi, \text{Id}) \lesssim ak^{-(1-s)} \).

2. \( \Theta(x, y) = (\theta(x, y), y) \) maps \( x \) almost to its right place, that is \( \theta(x, \psi(x, y)) - \phi_1(x, y) \ll 1 \). \( \Theta \) is defined via a construction similar to the construction (for \( s < 1/2 \)) in BBHM13; in order for it to work for \( s \in [1/2, 1) \), we need to regularize the flow (and therefore \( \theta(x, \psi(x, y)) \neq \phi_1(x, y) \)). We define \( \Theta \) in Section 3.3 show that \( \text{dist}_s(\Theta, \text{Id}) \lesssim k\lambda^{2-s} \delta^{-s} \), where \( \delta \ll \lambda \) is a regularization parameter to be determined. The main part of this section consists of proving bounds on \( \theta(x, \psi(x, y)) - \phi_1(x, y) \) and on the derivatives of \( \theta \).

3. In Section 3.4 we show that the error \( \Gamma = \Psi^{-1} \circ \Theta^{-1} \circ \Psi \circ \Phi_1^{-1} \) satisfies \( \text{dist}_s(\Gamma, \text{Id}) \lesssim \lambda^{1-s} \delta^{-(1-s)} \), by showing that the affine homotopy between \( \text{Id} \) and \( \Gamma \) is a path of small \( H^s \)-distance. This uses the bounds on \( \theta \) from Section 3.3.

Finally, we show that \( \alpha \) and \( \delta \) can be chosen such that, as \( k \to \infty \),

\[
\text{dist}_s(\Psi, \text{Id}) = o(1), \quad \text{dist}_s(\Theta, \text{Id}) = o(1), \quad \text{and} \quad \text{dist}_s(\Gamma, \text{Id}) = o(1),
\]

and then \( \text{dist}_s(\Phi_1, \text{Id}) = o(1) \) follows from Lemma 2.1.

Remark: Throughout this paper, we use big \( O \) and small \( o \) notations with respect to the limit \( k \to \infty \). We will also use notations such as \( \|I_j\| \approx k^{-1} \) above, meaning that there exist \( c_2 \geq c_1 > 0 \) such that \( c_1 k^{-1} \leq \|I_j\| \leq c_2 k^{-1} \). Finally, \( a \leq b \), means \( a \leq C b \) for some constant \( C \) (that can depend on the dimension \( n \) and the Sobolev exponent \( s \)).
Figure 1: A sketch of $\chi_k$. The solid part of the axis is $L_1$, where $\chi_k \equiv 1$. The dashed part of the axis is $L_2$.

### 3.1 Step I: Splitting into strips

Fix $k \in \mathbb{N}$. Define the following subintervals of $(0, 1)$:

$S_1^i := \left[ \frac{8i - 3}{k}, \frac{8i + 3}{k} \right], \quad L_1^i := \left[ \frac{8i - 2}{k}, \frac{8i + 2}{k} \right], \quad i \in \mathbb{Z},$

$S_2^i := \left[ \frac{8i + 1}{k}, \frac{8i + 7}{k} \right], \quad L_2^i := \left[ \frac{8i + 2}{k}, \frac{8i + 6}{k} \right], \quad i \in \mathbb{Z},$

and denote $S_j = \cup_i S_j^i \cap [0, 1], L_j = \cup_i L_j^i \cap [0, 1]$. Let $\chi : [-4, 4] \to [0, 1]$ be a smooth function satisfying $\text{supp} \chi \subset (-3, 3)$ and $\chi|_{[-2,2]} \equiv 1$. Extend $\chi$ periodically, and define $\chi_k(y) = \chi(ky)$ on $(0, 1)$. Note that $\text{supp} \chi_k \subset S_1, \chi_k|_{L_1} \equiv 1, \text{ and } |\chi_k'| \leq k$. See Figure 1.

Define $\zeta_1(x, y) = \zeta(x, y)\chi_k(y)$. Note that

$$\zeta_1|_{(0,1) \times L_1} = \zeta,$$

$$\text{supp}(\zeta_1) \subset (0, 1) \times S_1,$$  \hspace{1cm} (3.1)

and

$$0 \leq \zeta_1 \leq C, \quad -1 + C^{-1} < \partial_x \zeta_1 < C, \quad |\partial_y \zeta_1| < Ck,$$  \hspace{1cm} (3.2)

where $C$ is independent of $k$. Define

$$\Phi_1 = (\phi_1(x, y), y) = (x + \zeta_1(x, y), y), \quad \Phi_2 = \Phi \circ \Phi_1^{-1} = (\phi_2(x, y), y).$$

From (3.1)-(3.3), it immediately follows that we can write $\phi_2(x, y) = x + \zeta_2(x, y)$, with $\zeta_2$ satisfies properties (3.2) with $S_2$ and the bounds (3.3).

In the rest of this section we are going to prove that $\text{dist}_c(\Phi_1, \text{Id}) = o(1)$. This relies only on properties (3.2)-(3.3), hence the result also applies to $\Phi_2$, since $\zeta_2$ satisfies the same assumptions.

### 3.2 Step II: Squeezing the strips

**Lemma 3.2** Fix $\alpha \gg 1$. There exists a diffeomorphism $\Psi \in \text{Diff}_c(\mathbb{R}^2), \Psi(x, y) = (x, \psi(x, y))$, such that

$$\psi(x, y) = e^{-a} \left( y - \frac{8i}{k} \right) + \frac{8i}{k}, \quad (x, y) \in [0, 1] \times S_1^i \cap [0, 1],$$  \hspace{1cm} (3.4)
In other words, $\psi$ squeezes each intervals $S^i_1$ linearly around their midpoint by a factor of $e^{-\alpha}$, and has a small cost.

**Proof:** Let $u_1 \in C_c^\infty((-4,4))$, such that $u_1(y) = -y$ for $y \in [-3,3]$, and extend periodically. Let $\chi \in C_c^\infty(\mathbb{R}^3)$ such that $\chi \equiv 1$ on $[0,1]^2$. Define $u_k(x,y) := \frac{2}{k} u_1(ky) \chi(x,y)$.

Note that
\[
\|u_k\|_{L^2} \leq \alpha/k, \quad \|du_k\|_{L^2} \leq \alpha.
\]

Therefore, by Proposition 2.3 we have
\[
\|u_k\|_{H^{1-s}} \leq \frac{\alpha^{1-s}}{k^{1-s}} = \frac{\alpha}{k^{1-s}}.
\]

Let $\psi(t,x,y)$ be the solution of
\[
\partial_t \psi = u_k(x,\psi), \quad \psi(0,x,y) = y.
\]

Define $\psi(y) = \psi(1,y)$, and $\Psi(x,y) = (x, \psi(x,y))$. A direct calculation shows that for $(x,y) \in [0,1] \times [-3/k,3/k]$, $\psi(y) = ye^{-\alpha}$, so by periodicity and the fact that $\chi \equiv 1$ on $[0,1]^2$, $\psi$ satisfies (3.4).

The trajectory from $\text{Id}$ to $\Psi$ defined by $\Psi(t,x,y) = (x, \psi(t,x,y))$, together with the bound (3.6), imply (3.5). □

Note that in $[0,1]^2$, $\psi$ is independent of $x$. Therefore, slightly abusing notation, we write
\[
\Psi(x,y) = (x, \psi(y)), \quad \Psi^{-1}(x,y) = (x, \psi^{-1}(y)).
\]

We will later have $\alpha$ depend on $k$. Since eventually we want $\text{dist}_s(\Psi, \text{Id}) = o(1)$ when $k \to \infty$, (3.5) implies the bound
\[
\alpha \ll k^{1-s}.
\]

### 3.3 Step III: Flowing along the squeezed strips

Denote
\[
\lambda(x,k) = \frac{e^{-\alpha}}{k},
\]
and consider
\[
\Phi_1 \circ \Psi^{-1}(x,y) = (x + \zeta_1(x,\psi^{-1}(y)), \psi^{-1}(y)) =: (x + \tilde{\zeta}_1(x,y), \psi^{-1}(y)).
\]

Since $\zeta_1$ is supported inside $(0,1) \times S_1$, we have that $\tilde{\zeta}_1 = \zeta_1 \circ \Psi^{-1}$ is supported on $(0,1) \times \psi(S_1)$, that is, on $k$ strips of thickness $\approx \lambda$. Furthermore, from (3.3)-(3.4) we have
\[
\tilde{\zeta}_1 \geq 0, \quad -1 + C^{-1} < \partial_x \tilde{\zeta}_1 < C, \quad |\partial_y \tilde{\zeta}_1| < C \lambda^{-1}.
\]

We start by defining a path from $\text{Id}$ to
\[
\tilde{\Theta} := \Psi \circ \Phi_1 \circ \Psi^{-1}(x,y) = (x + \tilde{\zeta}_1(x,y), y),
\]

and
\[
\text{dist}_s(\Psi, \text{Id}) \leq ak^{-(1-s)}.
\]
using a slight variation of the construction of [BBHM13, Lemma 3.2] that proves that the $H^s$ geodesic distance is vanishing for $s < 1/2$. Let
\[ \tau_y(x) = x - \lambda \tilde{c}_1(x, y), \quad g_y = \tau_y^{-1}. \quad (3.9) \]
It is clear that $\tau_y$ is increasing for all small enough $\lambda$. We will henceforth restrict our attention to such $\lambda$, for which the definition of $g_y$ makes sense. We will also write $\tau(x, y)$ and $g(t, y)$ instead of $\tau_y(x)$ and $g_y(t)$. Define
\[ \tilde{\Theta}(t, x, y) = (\tilde{\theta}(t, x, y), y) \]
by
\[ \tilde{\theta}(t, x, y) := \begin{cases} 
  x & \text{if } t \leq \tau(x, y) \\
  x + (1 + \lambda)^{-1}(t - \tau(x, y)) & \text{if } \tau(x, y) \leq t \leq x + \tilde{c}_1(x) \\
  x + \tilde{c}_1(x, y) & \text{if } x + \tilde{c}_1(x) \leq t \leq 1.
\end{cases} \quad (3.10) \]
Note that $\tilde{\theta}$ solves
\[ \frac{\partial}{\partial t} \tilde{\theta}(t, x, y) = u(t, \tilde{\theta}(t, x, y), y), \quad \tilde{\theta}(0, x) = x, \]
where
\[ u(t, x, y) = u(t, x, y) := (1 + \lambda)^{-1} \mathbb{1}_{t < \tau(x, y)} = (1 + \lambda)^{-1} \mathbb{1}_{t \tau(x, y) < t < x}. \quad (3.11) \]
See Figure 2.
Lemma 3.3 The following bounds hold:

\[ g(t, y) = t + \lambda \tilde{\xi}_1(t, y) + O(\lambda^2), \quad g(t, y) = t \iff \tilde{\xi}_1(t, y) = 0. \]  \tag{3.12}

\[ \partial_1 g = 1 + \lambda \partial_1 \tilde{\xi}_1 + O(\lambda^2) = 1 + O(\lambda), \quad |\partial_2 g| < C. \]  \tag{3.13}

Proof: We fix \( y \) and write \( g(t) = g(t, y) \) and \( \tilde{\xi}_1(t) = \tilde{\xi}_1(t, y) \). Let \( \tilde{g}(t) = t + \lambda \tilde{\xi}_1(t) \), and let \( e(t) = g(t) - \tilde{g}(t) \). Then

\[ t = \tau(g(t)) = \tau(t + \lambda \tilde{\xi}_1(t) + e(t)) = t + \lambda \tilde{\xi}_1(t) + e(t) - \lambda \tilde{\xi}_1 \left( t + \lambda \tilde{\xi}_1(t) + e(t) \right). \]

Thus \( e = e(t) \) solves

\[ f(e; t) = e + \lambda \tilde{\xi}_1(t) - \lambda \tilde{\xi}_1 \left( t + \lambda \tilde{\xi}_1(t) + e(t) \right) = 0. \]

Since \( |f(0; t)| \leq \lambda^2 \| \partial_1 \tilde{\xi}_1 \|_\infty \| \tilde{\xi}_1 \|_\infty < C \lambda^2 \) for all \( t \) and \( \partial_1 f \geq 1 - \lambda \| \partial_1 \tilde{\xi}_1 \|_\infty \geq 1 - C \lambda \) (here we use (3.8)), the Intermediate Value Theorem implies that a unique \( e(t) \) such that \( f(e(t); t) = 0 \) and \( e(t) = O(\lambda^2) \). The second part of (3.12) is immediate from the definition of \( g \).

For proving (3.13), we use (3.8) and calculate

\[ \partial_1 g = \partial_1 \tau^{-1} = \frac{1}{\partial_1 \tau \circ g} = \frac{1}{1 - \lambda \partial_1 \tilde{\xi}_1 \circ g} = 1 + \lambda \partial_1 \tilde{\xi}_1 + O(\lambda^2), \]

and

\[ |\partial_2 g| = \left| \frac{\partial_2 \tau}{\partial_1 \tau} \right| = \left| \frac{\lambda \partial_2 \tilde{\xi}_1}{1 - \lambda \partial_1 \tilde{\xi}_1} \right| < C. \]

Since for every fixed \( x, \tilde{\xi}_1(x, \cdot) \) is supported on \( \approx k \) intervals of thickness \( \approx \lambda \), it follows from (3.11)-(3.12) that for every fixed \( t, u_t \) is supported on \( \approx k \) disjoint compact sets, each contained in a square of edge length \( \approx \lambda \), see Figure 3.

Since \( u_t \notin H^s \) for \( s \geq 1/2 \), we need to regularize. To do this, fix \( \delta \ll \lambda \) (to be determined) and define

\[ u_{\delta,t}(x, y) = u_\delta(t, x, y) := \int \limits_{\mathbb{R}} u(t, x - x', y) \eta_\delta(x') dx' = \frac{1}{1 + \delta} \int_{x - g(t, y)}^{x - t} \eta_\delta(x') dx' \]  \tag{3.14}

for \( \eta_\delta \in C_c^\infty(\mathbb{R}) \) such that

\[ \eta_\delta \geq 0, \quad \int_{-\infty}^0 \eta_\delta = \int_0^\infty \eta_\delta = \frac{1}{2}, \quad \text{supp}(\eta_\delta) \subset [-\delta, \delta], \quad \| \eta_\delta \|_\infty \leq \frac{C}{\delta}. \]

Lemma 3.4 For a fixed \( t, u_{\delta,t}(x, y) \in W^{1,\infty}(\mathbb{R}^2) \), and

\[ \| du_{\delta,t} \|_\infty \leq \frac{1}{\delta}. \]  \tag{3.15}

Moreover,

\[ \| u_{\delta,t} \|_{L^p}^2 \leq \frac{k \lambda^{2-s}}{\delta^s}. \]  \tag{3.16}
Proof: $|\partial_t u_\delta| < C/\delta$ follows from the definition of $u_\delta$ and the bounds on $\eta_\delta$. We now show that $u_\delta$ is also Lipschitz with respect to the $y$ variable. Indeed, note that

$$
|u(t,x,y'+h)-u(t,x,y')| = (1+\lambda)^{-1} \mathbb{1}_{g(t,y)<x<g(t,y+h)},
$$

if $g(t,y+h) > g(t,y)$, and similarly if not. By (3.13),

$$
|g(t,y+h)-g(t,y)| \leq |h|\|\partial^2 g\|_{\infty} \leq C|h|
$$

and therefore we have

$$
\|u(t,\cdot,y'+h)-u(t,\cdot,y')\|_1 \leq (1+\lambda)^{-1}C|h| \leq |h|.
$$

Finally,

$$
|u_\delta(t,x,y'+h)-u_\delta(t,x,y')| \leq \|\eta_\delta\|_{\infty} \|u(t,\cdot,y'+h)-u(t,\cdot,y')\|_1 \leq \frac{|h|}{\delta},
$$

which completes the proof of (3.15).

Now, similar to $u_t$, $u_\delta_t$ is supported on $\approx k$ disjoint compact sets, each contained in a square of edge length $\approx \lambda$. Since $u_t$ is an indicator function, $du_t$ is supported on a $\delta$-neighborhood of the boundary of supp $u_t$. Since $|\partial^2 g| \leq C$ (see (3.13)), it follows that $du_\delta_t$ is supported on $\approx k$ sets of area of $\approx \delta\lambda$ (see Figure 3).

Since $|u_\delta|_{\infty} < 1$, and $u_\delta_t$ is supported on a set of measure $\approx k\lambda^2$, we have

$$
\|u_\delta\|^2 \lesssim k\lambda^2.
$$
Since $|du_{\delta,t}| \leq C/\delta$, and $du_{\delta,t}$ is supported on a set of measure $\approx k\lambda \delta$,

$$\|du_{\delta,t}\|_2^2 \lesssim \frac{k\lambda}{\delta}. \quad (3.16)$$

follows from these bounds and Proposition 2.3.

Since we eventually want $u_{t,\lambda}$ to have a small $H^s$ norm, we will henceforth assume that $\delta$ satisfies

$$k\lambda^{2-s} \ll \delta^s \ll k^{-s/(1-s)} \lambda^s, \quad (3.17)$$

where the upper-bound assumption (which is more restrictive than the natural $\delta \ll \lambda$) will be needed later. In particular, note that these assumptions put some restrictions on the possible choices of $\lambda = e^{-\alpha}/k$, in addition to (3.7). We will give concrete choices of $\alpha$ and $\delta$ that satisfy these bounds in the end of the proof in Section 3.4.

Let $\theta(t, x, y)$ be the solution of

$$\frac{\partial}{\partial t} \theta(t, x, y) = u_{\delta}(t, \theta(t, x, y), y), \quad \theta_{\delta}(0, x) = x \quad (3.18)$$

and define $\theta(x, y) = \theta(1, x, y)$. Define $\Theta \in \text{Diff}_c(\mathbb{R}^2)$ by

$$\Theta(x, y) = (\theta(x, y), y). \quad (3.19)$$

It follows immediately from (3.16) that

$$\text{dist}_s(\text{Id}, \Theta) \lesssim \frac{k^{1/2}(2-s)/2}{\delta^{s/2}} \quad (3.20)$$

The following Lemma states that the amount $\Theta$ “misses” the target $\tilde{\Theta}$ because of the mollification is small:

**Lemma 3.5** $\text{supp}(\theta(x, y) - x)$ is a subset of a $\delta$-thickening in the $x$ direction of $\text{supp}(\tilde{\zeta}_1)$, that is

$$\text{supp}(\theta(x, y) - x) \subset \{(x, y) : \exists (x', y) \in \text{supp}(\tilde{\zeta}_1), |x - x'| < \delta\}. \quad (3.21)$$

In particular, for small enough $\delta$, $\text{supp}(\theta(x, y) - x) \subset (0, 1)^2$. Moreover,

$$|\theta(x, y) - (x + \tilde{\zeta}_1(x, y))| \leq \frac{3\delta}{\lambda} \quad (3.22)$$

**Proof:** Throughout this proof $y$ is fixed and does not play a role, and we will omit it for notational brevity. Conclusion (3.21) follows immediately from the definition of $\theta$. We now prove (3.22). Define

$$u_{\delta}^- = (1 + \lambda)^{-1} 1_{\{u_{\delta} = (1+\lambda)^{-1}\}} = (1 + \lambda)^{-1} 1_{t+\delta < g(t) - \delta},$$

$$u_{\delta}^+ = (1 + \lambda)^{-1} \frac{1}{2} (1_{\text{supp } u} + 1_{\text{supp } u_{\delta}}) = (1 + \lambda)^{-1} \left( 1_{t < g(t)} + \frac{1}{2} 1_{\text{supp } u_{\delta} \setminus \text{supp } u} \right)$$

and let $\theta^\pm(t, x)$ solve

$$\frac{\partial}{\partial t} \theta^\pm(t, x) = u_{\delta}^\pm(t, \theta^\pm(t, x)), \quad \theta^\pm(0, x) = x.$$
Figure 4: A sketch of the flow $\theta^+$ along $u_0^+$. The dark grey area is supp $u$, where $u_0^+ = (1 + \lambda)^{-1}$. The light grey area is supp $u_0 \setminus$ supp $u$, which is at most of width $\delta$; in this region $u_0^+ = \frac{1}{2}(1 + \lambda)^{-1}$.

and let $\theta^\pm(x) := \theta^\pm(1, x)$.

It is clear that

$$u_0^- \leq u_0 \leq u_0^+$$

pointwise. It follows that $\theta^-(t, x) \leq \theta(t, x) \leq \theta^+(t, x)$ for all $t \geq 0$ and all $x$, and in particular $\theta^-(x) \leq \theta(x) \leq \theta^+(x)$. See Figure 4

First consider $\theta^+(t, x)$. Note that $\theta^+(t, x) = x$ for $t \leq t_1$, where is the first time such that $(t_1, x) \in$ supp $u_0$. Since supp $\eta \subset [-\delta, \delta]$ we have

$$t_1 \geq \tau(x - \delta).$$

Since $\partial_1 \tau = 1 + O(\lambda)$ (see (3.8)-(3.9)), it follows that $t_1 \geq \tau(x) - 2\delta$. From $t_1$, until time $t_2$ defined by

$$g(t_2) = \theta^+(t_2, x),$$

i.e. the first time such that $(t_2, \theta^+(t_2, x)) \in$ supp $u$, we have $\theta^+(t, x) < x + \frac{1}{2}(t - t_1)$ (note that for certain values of $x$, $(t, \theta^+(t, x)) \notin$ supp $u$ for any $t$. In this case the analysis is simpler). Using this inequality, (3.13) and the bound on $t_1$, it follows that $t_2 - t_1 \leq 5\delta$. Indeed,

$$x + \frac{1}{2}(t_2 - t_1) > \theta^+(t_2, x) = g(t_2) > g(\tau(x)) + (1 - C\lambda)(t_2 - \tau(x))$$

and since $g(\tau(x)) = x$, we see that $\frac{1}{2}(t_2 - t_1) > (1 - C\lambda)(t_2 - t_1 - 2\delta)$, from which the claim follows. Therefore $\theta^+(t_2, x) < x + 3\delta$. Until the time $t_3$ when $\theta^+(t, x)$ leaves supp $u$, $\theta^+$ flows
according to the flow of $u$ with initial condition $\theta^+(t_2, x)$. Therefore,

$$\theta^+(t_3, x) = \theta^+(t_2, x) + \tilde{\zeta}_1(\theta^+(t_2, x)) < x + \tilde{\zeta}_1(x) + C\delta,$$

where we used (3.8) again. By the same arguments as the time interval $[t_1, t_2]$, it follows that for $t > t_3$, $\theta^+(t, x)$ increases by less than $\delta$. Therefore we obtain the upper bound

$$\theta(x) \leq \theta^+(x) < x + \tilde{\zeta}_1(x) + C\delta,$$  \hspace{1cm} (3.23)

for an appropriate constant $C$.

We now consider $u_\delta^-$ and $\theta^-(t, x)$. Note that

$$u_\delta^-(t, x) = (1 + \lambda)^{-1} \mathbb{1}_{\tau(x, \delta) < t < x - \delta} > (1 + \lambda)^{-1} \mathbb{1}_{\tau(x, \delta) + 2\delta < t < x - \delta},$$

where we used $\partial_1 \tau = 1 + O(\lambda)$ in the inequality. Defining $t' = t + \delta$, we have

$$u_\delta^-(t', x) \geq v_\delta^-(t', x) := (1 + \lambda)^{-1} \mathbb{1}_{\max\{\tau(x, 3\delta, x) < t' < x \}).$$

By the definition (3.9) of $\tau$

$$\tau(x) + 3\delta = x - \lambda \tilde{\zeta}_1(x) + 3\delta = x - \lambda \left(\tilde{\zeta}_1(x) - \frac{3\delta}{\lambda}\right).$$

It follows that the flow by $v_\delta^-(t, x)$, that is the solution $\tilde{\theta}^-$ of

$$\frac{\partial}{\partial t} \tilde{\theta}^-(t, x) = v_\delta^-(t, \tilde{\theta}^-(t, x)) \quad \tilde{\theta}^-(0, x) = x,$$

satisfies

$$\tilde{\theta}^-(1, x) = \max \left\{ x + \tilde{\zeta}_1(x) - \frac{3\delta}{\lambda}, x \right\}.$$

Moreover, for $\delta$ small enough (depending only on $\zeta$), $\tilde{\theta}^-(1 - \delta, x) = \tilde{\theta}^-(1, x)$. By (3.24), it follows that

$$\theta(x) \geq \theta^-(1, x) \geq \tilde{\theta}^-(1 - \delta, x) \geq x + \tilde{\zeta}_1(x) - \frac{3\delta}{\lambda}.$$  \hspace{1cm} (3.25)

(3.23) and (3.25) imply (3.22).

Next, we now prove bounds on the derivatives of $\theta$.

**Lemma 3.6** There exists $C \geq 1$, depending only on $\zeta$, such that

$$C^{-1} \leq \partial_x \theta \leq C \quad \text{for all} \ (x, y).$$  \hspace{1cm} (3.26)

**Proof:** As in the proof of Lemma 3.5, we will omit $y$ for notational brevity, and because it does not play any role. Recall that $\partial_2 \theta(t, x) = u_\delta(t, \theta)$, and consider the Eulerian version of this flow, that is the equation

$$\partial_t w(t, x) + u_\delta(t, x) \partial_x w(t, x) = 0$$  \hspace{1cm} (3.27)

with initial data

$$w(0, x) = x.$$  \hspace{1cm} (3.28)
If \( w \) is a solution then
\[
\frac{d}{dt} w(t, \theta(t, x)) = \partial_t w(t, \theta) \partial_t \theta + \partial_t w(t, \theta) = 0,
\]
using the ODE for \( \theta \) and the PDE for \( w \). The initial data then imply that \( w(t, \theta(t, x)) = x \) for all \( t \), and hence that
\[
w(t, \cdot) = \theta(t, \cdot)^{-1}.
\]
Next, define
\[q = \partial_t w + \partial_x w.\]
Since \( u_\delta(t, x) = 0 \) when \( t \) is close to 0 or 1, we have that \( \partial_t w = 0 \) for such values of \( t \). In particular, \( q(0, \cdot) = 1 \) and \( q(1, \cdot) = \partial_x w(1, \cdot) = \partial_t \theta(1, \cdot)^{-1} \), which is the quantity we need to estimate.

We use \( q \) and not \( \partial_x w \) directly since it will allow us to exploit the fact, reflected in the smallness of \( (\partial_t + \partial_x)u_\delta \), that the coefficients in (3.27) are nearly translation-invariant in the \( \partial_t + \partial_x \) direction. We compute
\[
\partial_t q = \partial_t (\partial_t w + \partial_x w) = (\partial_t + \partial_x)\partial_t w = -(\partial_t + \partial_x)(u_\delta \partial_x w) = -u_\delta \partial_x q - (\partial_t u_\delta + \partial_x u_\delta) \partial_x w.
\]
We further deduce from (3.27) that
\[
\partial_x w = q + u_\delta \partial_x w, \quad \text{and thus} \quad \partial_x w = \frac{q}{1 - u_\delta},
\]
so we can rewrite the above equation as
\[
\partial_t q = -u_\delta \partial_x q - \frac{\partial_t u_\delta + \partial_x u_\delta}{1 - u_\delta} q.
\]
It follows that
\[
\frac{d}{dt} q(t, \theta(t, x)) = -\frac{\partial_t u_\delta + \partial_x u_\delta}{1 - u_\delta} \big(t, \theta(t, x)\big) q(t, \theta(t, x)). \tag{3.29}
\]
Therefore, if we obtain a bound
\[
\int_0^1 \left| \frac{\partial_t u_\delta + \partial_x u_\delta}{1 - u_\delta} \big(t, \theta(t, x)\big) \right| dt < C, \tag{3.30}
\]
for some \( C \) independent of \( x \) (and \( y \)), we obtain (3.26) by Gronwall’s inequality.

From the definition (3.14) of \( u_\delta \), we have
\[
\partial_x u_\delta(t, x) = \frac{1}{1 + \lambda} \left[ \eta_\delta(x - t) - \eta_\delta(x - g(t)) \right], \tag{3.31}
\]
\[
\partial_t u_\delta(t, x) = \frac{1}{1 + \lambda} \left[ -\eta_\delta(x - t) + g'(t) \eta_\delta(x - g(t)) \right], \tag{3.32}
\]
and therefore, using (3.13), we have
\[
|\partial_t u_\delta + \partial_x u_\delta| = \frac{1}{1 + \lambda} \eta_\delta(x - g(t)) \left| g'(t) - 1 \right| \leq \frac{C \lambda}{1 + \lambda} \eta_\delta(x - g(t)). \tag{3.33}
\]
Because of (3.29) and (3.33), we want to estimate $\frac{\eta_0(x-g(t))}{1-u_0(t,x)}$. We have

$$1 - u_0(t,x) = 1 - \frac{1}{1 + \lambda} \int_{x-g(t)}^{x-t} \eta_0(x') dx'$$

$$\geq 1 - \frac{1}{1 + \lambda} \int_{x-g(t)}^{\infty} \eta_0(x') dx'$$

$$= 1 - \frac{1}{1 + \lambda} \mu_0(x - g(t)), \quad \text{for} \quad \mu_0(x) := \int_{x}^{\infty} \eta_0(x') dx',$$

and therefore

$$\frac{\eta_0(x - g(t))}{1 - u_0(t,x)} \leq \frac{(1 + \lambda)\eta_0(x - g(t))}{1 + \lambda - \mu_0(x - g(t))} = \frac{(1 + \lambda)\mu_0'(x - g(t))}{1 + \lambda - \mu_0(x - g(t))}.$$ 

It follows that

$$\int_{0}^{1} \left| \frac{\partial u_0 + \partial_x u_0 (t, \theta(t,x))}{1 - u_0} \right| dt \leq C \lambda \int_{0}^{1} \frac{-\mu_0'(\theta(t,x) - g(t))}{1 + \lambda - \mu_0(\theta(t,x) - g(t))} dt. \quad (3.34)$$ 

For the following computation, $x$ is fixed. We wish to rewrite the integral in terms of the variable

$$\alpha = \alpha(t) = \mu_0(\theta(t,x) - g(t)),$$

which increases from 0 to 1 as $t$ goes from 0 to 1 for $\delta, \lambda$ sufficiently small. To estimate $\alpha'(t)$, note that by the definition of $\theta$, we have

$$\partial_t \theta(t,x) = u_0(t, \theta(t,x)) = \frac{1}{1 + \lambda} \int_{\theta(t,x) - g(t)}^{\theta(t,x) - t} \eta_0(x') dx' \leq \frac{1}{1 + \lambda} \int_{\theta(t,x) - g(t)}^{\infty} \eta_0(x') dx'$$

$$= \frac{\alpha(t)}{1 + \lambda}.$$ 

Since $g' \geq 1 - c\lambda$ for some $c < 1$, depending only on $\zeta$, it follows that

$$\partial_t (\theta(t,x) - g(t)) \leq \frac{\alpha(t)}{1 + \lambda} - 1 + c\lambda = \frac{\alpha(t) - (1 + \lambda)(1 - c\lambda)}{1 + \lambda}.$$ 

This is always negative for small enough $\lambda$, as $0 \leq \alpha \leq 1$ and $c < 1$. Thus

$$-\mu_0'(\theta(t,x) - g(t)) = \frac{\alpha'(t)}{\alpha}(\theta(t,x) - g(t)) \leq \frac{(1 + \lambda)\alpha'(t)(\theta(t,x) - g(t))}{(1 + \lambda)(1 - c\lambda) - \alpha(t)}.$$ 

So we can change variables in (3.34) to find that

$$\int_{0}^{1} \left| \frac{\partial u_0 + \partial_x u_0 (t, \theta(t,x))}{1 - u_0} \right| dt \leq C \lambda \int_{0}^{1} \frac{1}{1 + \lambda - \alpha} \frac{1}{(1 + \lambda)(1 - c\lambda) - \alpha} d\alpha.$$ 

For $\lambda < \frac{1 - c}{2\lambda}$, the integrand on the right is bounded by $(1 + \frac{1}{2}(1 - c)\lambda - \alpha)^{-2}$, so we integrate to conclude that

$$\int_{0}^{1} \left| \frac{\partial u_0 + \partial_x u_0 (t, \theta(t,x))}{1 - u_0} \right| dt \leq C \lambda (\frac{1}{2}(1 - c)\lambda)^{-1} \leq C.$$ 

We thus obtain (3.30), which completes the proof. 

\[\blacksquare\]
Lemma 3.7 For every \( \lambda > 0 \) small enough, there exists a choice of mollifier \( \eta_\delta \) in the definition (3.14) such that
\[
|\partial_y \theta| \leq C\lambda^{-1} \quad \text{for all} \ (x, y),
\]
where \( C > 0 \) depends only on \( \zeta \).

Proof: Fix \( h \in \mathbb{R}, |h| \ll \delta \lambda \), and consider \( \theta(t, x, y) \) and \( \theta(t, x, y + h) \). By Lemma 3.3 we have that
\[
|\tau(t, y + h) - \tau(t, y)| < c|h|,
\]
for some \( c > 0 \). In particular,
\[
u(t - c|h|, x, y + h) = (1 + \lambda)^{-1}1_{|\tau(x, y + h)|<c|h|<x} 
\]
\[
\leq (1 + \lambda)^{-1}1_{|\tau(x, y)|<c|h|<x} = (1 + \lambda)^{-1}1_{|\tau(x, y)|<c|h|+h} 
\]
\[
= (1 + \lambda)^{-1}1_{1-c|h|<x}(t, y) =: u^h(t, x, y).
\]
Therefore \( \nu_\delta(t, x, y + h) \leq u^h_\delta(t + c|h|, x, y) \), where \( u^h_\delta \) is the mollification of \( u^h \) as in (3.14).

Define \( \theta^h(t, x, y) \) by
\[
\frac{\partial}{\partial t} \theta^h(t, x, y) = u^h_\delta(t, \theta(t, x, y), y), \quad \theta_\delta(0, x) = x.
\]
It follows that
\[
\theta(t - c|h|, x, y + h) \leq \theta^h(t, x, y),
\]
and since for \( h \) small enough (independent of \( x \) and \( y \) ), \( \theta(1, x, y + h) = \theta(1 - c|h|, x, y + h) \), we have
\[
\theta(1, x, y + h) \leq \theta^h(1, x, y).
\]

We now compare \( \theta^h(t, x, y) \) and \( \theta(t, x, y) \) and show that
\[
\theta(1, x, y + h) - \theta(1, x, y) \leq \theta^h(1, x, y) - \theta(1, x, y) \leq \frac{|h|}{\lambda}.
\]
(3.36)

By symmetry it also follows that
\[
\theta(1, x, y) - \theta(1, x, y + h) \leq \frac{|h|}{\lambda},
\]
which completes the proof.

It remains to prove the righthand side inequality in (3.36). In order to simplify notation, we will henceforth write \( \theta(t) = \theta(t, x, y), g(t) = g(t, y) \) and so on.

For this, it is convenient to use a smooth mollifier \( \eta_\delta \) with support in \([-\delta, \delta]\) such that
\[
0 \leq \eta_\delta(x) \leq \frac{1 + \lambda}{2\delta}.
\]

This is necessarily very close to the normalized characteristic function of the interval \([-\delta, \delta]\) in \( L^p \) for every \( p < \infty \). By the definition of \( \theta^h(t) \), it follows (see Figure 5) that
\[
\theta^h(t) = \theta(t).
\]
for every \( t < t_0 \), where \( t_0 \) is defined by
\[
\theta(t_0) = t_0 + \delta.
\]

When \( \theta(t) - t \geq -\delta \), we have
\[
\frac{d}{dt} \theta = u(t, \theta) = \frac{1}{1 + \lambda} \int_{\theta-g(t)}^{\theta-t} \eta(x')dx' \leq \frac{1}{1 + \lambda} \int_{-\infty}^{\theta-t} \eta(x')dx' \leq \min \left\{ \frac{1}{1 + \lambda} (\theta - t + \delta), \frac{1}{1 + \lambda} \right\},
\]
and when \( \theta(t) - t \leq -\delta \) we have \( \frac{d\theta}{dt} = 0 \). Let \( \alpha(t) = \theta(t) - t \). It follows that
\[
\frac{d\alpha}{dt} \leq -\frac{\lambda}{1 + \lambda} + \min \left\{ \frac{1}{1 + \lambda} (\alpha - \delta - \frac{1 - \lambda}{1 + \lambda}), 0 \right\} \quad \text{as long as } \alpha(t) \geq -\delta,
\]
and \( \frac{d\alpha}{dt} = -1 \) when \( \alpha(t) \leq -\delta \). If we write \( \alpha_0(t) \) to denote the function solving the above ODE (with \( \leq \) replaced by \( = \)) with initial data \( \alpha_0(t_0) = \delta \), then \( \alpha(t) \leq \alpha_0(t) \) for \( t \geq t_0 \). This leads to
\[
\alpha(t) \leq \begin{cases} 
\delta - \frac{\lambda}{1 + \lambda} (t - t_0) & \text{if } t_0 \leq t \leq t_a = t_0 + 2\delta \\
\delta - \delta \frac{2\lambda}{1 + \lambda} \exp \left( \frac{t - t_0}{2\delta} \right) & \text{if } t_a \leq t
\end{cases}
\]
as long as \( \alpha(t) \geq -\delta \).

We now define \( t_1 \) to be the unique time such that \( \alpha(t_1) = -\delta \), and similarly \( t_2 \) such that \( \alpha(t_2) = -2\delta \) (see Figure 5). We deduce from the above that
\[
t_1 - t_0 \leq 2\delta \left( 1 + \log \left( \frac{1 + \lambda}{\lambda} \right) \right), \quad t_2 - t_1 = \delta.
\]

(3.37)
Next we estimate $\theta^h(t_2) - \theta(t_2)$. First note that

$$0 \leq u^h_\delta(t, x) - u_\delta(t, x) = \frac{1}{1 + \lambda} \int_{x-t}^{x-t+h|t|} \eta_\delta(x') dx' \leq \frac{|h|}{2\delta}.$$ 

We can similarly estimate $u_\delta(t, x') - u_\delta(t, x)$, to find that

$$\frac{d}{dt}(\theta^h - \theta) = [u^h_\delta(t, \theta^h) - u_\delta(t, \theta^h)] + [u_\delta(t, \theta^h) - u_\delta(t, \theta)] \leq \frac{|h|}{2\delta} + \frac{1}{2\delta}(\theta^h - \theta).$$

(We have implicitly used the fact that $\theta^h(t) \geq \theta(t)$ for all $t$). Thus Grönwall’s inequality implies that for $t > t_0$,

$$\theta^h(t) - \theta(t) \leq c|h|\left[ \exp\left( \frac{t - t_0}{2\delta} \right) - 1 \right].$$

In particular it follows from (3.37) that

$$\theta^h(t_2) - \theta(t_2) \leq c|h|\left[ \exp\left( \frac{3}{2} + \log\left( \frac{1 + \lambda}{\lambda} \right) \right) - 1 \right] \leq \frac{|h|}{\lambda}.$$ 

Thus

$$\theta^h(t_2) - t_2 = \theta^h(t_2) - \theta(t_2) + \alpha(t_2) \leq -2\delta + c\frac{|h|}{\lambda} < -\delta$$

(3.38)

for $h$ small enough. Since $\theta^h(t) - t$ is a decreasing function, this inequality continues to hold after time $t_2$. It then follows from the definitions that $u^h_\delta(t, \theta^h(t)) = u_\delta(t, \theta(t)) = 0$ for $t \geq t_2$, and therefore

$$\theta^h(1) - \theta(1) = \theta^h(t_2) - \theta(t_2) \approx \frac{|h|}{\lambda},$$

which proves (3.36) and completes the proof.

\section*{3.4 Step IV: Error correction – affine homotopy}

\textbf{Corollary 3.8} The diffeomorphism $\Psi^{-1} \circ \Theta \circ \Psi$ is of the form

$$\Psi^{-1} \circ \Theta \circ \Psi = (x + \sigma(x, y), y),$$

(3.39)

where $\sigma(x, y) \geq 0$ is supported on $(0, 1) \times S_1$ and satisfies

$$|\sigma(x, y) - \zeta_1(x, y)| \leq \frac{\delta}{\lambda}, \quad -1 + C^{-1} < \partial_x \sigma < C, \quad |\partial_y \sigma| \leq k.$$ (3.40)

\textbf{Proof:} Conclusion (3.39) is immediate from the definitions of $\Psi$ and $\Theta$. We see from (3.21) that supp $\sigma$ is a subset of a $\delta$-thickening in the $x$-direction of supp($\zeta_1$). Therefore, (3.2) implies supp($\sigma$) $\subset (0, 1) \times S_1$ for small enough $\delta$. The first bound in (3.40) follows from Lemma 3.5, the second from Lemma 3.6, and the third from Lemma 3.7 using the fact that $\Psi$ is linearly squeezing strips on which $\theta$ is supported by a factor of $e^{-\alpha} = k\lambda$. \hfill \square
Corollary 3.9 \textit{The diffeomorphism }$\Gamma = \Psi^{-1} \circ \Theta \circ \Psi \circ \Phi^{-1}$\textit{ is of the form}

$$\Gamma = (\gamma(x, y), y) = (x + \xi(x, y), y), \quad (3.41)$$

\textit{where }$\xi(x, y) \geq 0$\textit{ is supported on }$(0, 1) \times S_1$\textit{ and satisfies}

$$|\xi(x, y)| \leq \frac{\delta}{\lambda}, \quad -1 + C^{-1} < \partial_x \xi < C, \quad |\partial_y \xi| \leq k. \quad (3.42)$$

\textbf{Proof:} This is immediate from the previous corollary, the definition of $\Phi_1$ and the bounds (3.3). \hfill \blacksquare

Lemma 3.10

$$\text{dist}_s(\Gamma, \text{Id}) \lesssim \frac{\delta^{1-s}}{\lambda^{1-s}} k^s. \quad (3.43)$$

\textbf{Proof:} Consider an affine homotopy $\Gamma_t$ from $\text{Id}$ to $\Gamma$, that is,

$$\Gamma_t(x, y) = (x + t\xi(x, y), y) =: (\gamma_{t,y}(x), y)$$

We then have $\partial_t \Gamma_t = u_t(\Gamma_t)$, where

$$u_t(x, y) = (\xi(\Gamma_t^{-1}(x, y)), 0) = (\xi(\gamma_{t,y}^{-1}(x), y), 0).$$

Note that $u_t$ is supported on a subset of the unit square, because $\xi$ is supported on a subset of the unit square and $\Gamma$ is a diffeomorphism of the unit square. Since $|\xi| \leq \delta \lambda^{-1}$, we have

$$\|u_t\|_{L^2} \lesssim \frac{\delta}{\lambda}. \quad (3.44)$$

Next, we have

$$\partial_x u_t(x, y) = \partial_x \xi \partial_x \gamma_{t,y}^{-1}(x), \quad \partial_y u_t(x, y) = \partial_x \xi \partial_y \gamma_{t,y}^{-1}(x) + \partial_y \xi.$$

Since, by (3.42), $-1 + C^{-1} < \partial_x \xi < C$, we obtain that $|\partial_x \gamma_{t,y}^{-1}| = |1 + t \partial_x \xi|^{-1} < C$ and therefore $|\partial_x u_t| < C$. Next, using (3.42) again, we have

$$\left| \frac{\partial_y \gamma_{t,y}^{-1}(x)}{\partial_x \gamma_{t,y}} \right| \lesssim k,$$

and therefore $|\partial_y u_t| \lesssim k$. We conclude that

$$\|u_t\|_{H^1} \lesssim k. \quad (3.45)$$

Using Proposition 2.3, (3.44)-(3.45) imply (3.43). \hfill \blacksquare

We conclude now the proof of Theorem 3.1. We showed that

$$\Phi_1 = \Gamma^{-1} \circ \Psi^{-1} \circ \Theta \circ \Psi, \quad \Gamma, \Theta, \Psi \in \text{Diff}_c(\mathbb{R}^2),$$

where (following Lemma 3.2, (3.20) and Lemma 3.10)

$$\text{dist}_s(\Psi, \text{Id}) \lesssim ak^{-(1-s)}, \quad \text{dist}_s(\Theta, \text{Id}) \lesssim \frac{k^{1/2} \lambda^{(2-s)/2}}{\delta^{s/2}}, \quad \text{dist}_s(\Gamma, \text{Id}) \lesssim \frac{\delta^{1-s}}{\lambda^{1-s}} k^s, \quad \lambda = e^{-\alpha} k.$$
If we choose, say
\[ \alpha = (\log k)^2, \quad \lambda = \frac{1}{k^{1+\log k}}, \quad \delta = \frac{1}{k^{\log k + \sqrt{\log k}}}, \]
we have, for any \( s < 1 \),
\[
\dist_s(\Theta, \text{Id}) \lesssim (\log k)^2 k^{-(1-s)} = o(1),
\]
\[
\dist_s(\phi, \text{Id}) \lesssim k^{-(1-s)\log k + \frac{1}{2}s} \sqrt{\log k - \frac{1}{2}} = o(1),
\]
\[
\dist_s(\Gamma, \text{Id}) \lesssim k^{1-(1-s)} \sqrt{\log k} = o(1),
\]
and therefore \( \dist_s(\Phi, \text{Id}) = o(1) \), which completes the proof.

Remark: Since we choose \( \alpha \) and \( \delta \) in an \( s \)-independent way, we constructed a sequence of paths from \( \text{Id} \) to \( \Phi \) that are of asymptotically vanishing \( H^s \)-cost for any \( s < 1 \). It follows that by choosing appropriate sequences of exponents \( s_n \nearrow 1 \) and constants \( c_n \searrow 0 \), we have
\[ \dist_{H^{s_n}}(\Phi, \text{Id}) = 0, \]
where the \( H^{s_n} \)-norm is defined by
\[ \|f\|_{H^{s_n}} := \sum_{n=1}^{\infty} c_n \|f\|_{H^{s_n}}. \]

4 Higher-dimensional construction

In this section we present a simpler construction in \( \mathbb{R}^n \) for \( n \geq 3 \). Since we often want to split \( \mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1} \), it is convenient to write \( m = n - 1 \).

**Theorem 4.1** Let \( n \geq 3 \), and denote by \( (x, y) \) the coordinates on \( \mathbb{R}^n \), where \( x \in \mathbb{R} \) and \( y \in \mathbb{R}^m \). Let \( \zeta \in C^\infty((0, 1)^m) \) satisfying \( \zeta \geq 0, \partial_1 \zeta > -1 \). Denote \( \phi(x, y) = x + \zeta(x, y) \). Define \( \Phi \in \text{Diff}_c(\mathbb{R}^{1+m}) \) by \( \Phi(x, y) = (\phi(x, y), y) \). Then \( \dist_s(\Phi, \text{Id}) = 0 \) for every \( s \in [0, 1) \).

While in principle one can adjust the construction from the two-dimensional case to this setting, we can take advantage of the fact of the higher dimensionality to make a simpler construction, as outlined below: First, in Section 4.1 we decompose \( \Phi \) as follows:
\[ \Phi = \Phi_2 \circ \ldots \circ \Phi_2 \circ \Phi_1, \quad \Phi_i = (\phi_i(x, y), y) = (x + \zeta_i(x, y), y) \in \text{Diff}_c(\mathbb{R}^{1+m}), \]
where \( \zeta_i \) is supported on the union of \( \approx k^m \) “tubes” \( (0, 1) \times I_j \), where \( I_j \) are \( m \)-dimensional cubes of edge length \( \approx k \). This is a generalization of the construction in Section 3.1. In the rest of Section 4 we show that \( \dist_s(\Phi_1, \text{Id}) = o(1) \) as \( k \to \infty \), and the same holds for all the other \( \Phi_i \)s. Since \( k \) is arbitrary, the conclusion \( \dist_s(\Phi, \text{Id}) = 0 \) follows by Lemma 2.1.

In order to prove \( \dist_s(\Phi_1, \text{Id}) = o(1) \), we decompose \( \Phi_1 \) as
\[ \Phi_1 = \Psi^{-1} \circ \Gamma \circ \Psi, \quad \Psi, \Gamma \in \text{Diff}_c(\mathbb{R}^{1+m}), \]
where
1. \( \Psi(x, y) = (x, \psi(x, y)) \) squeezes the cubes \( I_j \) on which \( \Phi_1 \) is supported by a factor of \( k^{\log k} \).

In Section 4.2, we define \( \Psi(x, y) \) and show that \( \text{dist}_s(\Phi, \text{Id}) \lesssim (\log k)^2 k^{-(1-s)} = o(1) \). This is analogous to Section 3.2 with \( a = (\log k)^2 \).

2. \( \Gamma = \Psi \circ \Phi_1 \circ \Psi^{-1} \). Unlike in the two-dimensional case, we do not have to construct a complicated flow along the strips (as in Section 3.3, which is the main part of the proof). Instead, in Section 4.3, we show that the affine homotopy between \( \text{Id} \) and \( \Gamma \) is a path of small \( H^s \) distance, and therefore \( \text{dist}_s(\Gamma, \text{Id}) \lesssim k^{(m/2-s) \log k} = o(1) \).

It then follows from Lemma 2.1 that \( \text{dist}_s(\Phi_1, \text{Id}) = o(1) \).

4.1 Step I: Splitting into strips

Fix \( k \in \mathbb{N} \), and consider the lattice \( \frac{1}{k} \mathbb{Z}^m \subset \mathbb{R}^m \). We partition \( \mathbb{Z}^m \) into \( 2^m \) lattices:

\[
2 \mathbb{Z}^m, 2 \mathbb{Z}^m + e_1, \ldots, 2 \mathbb{Z}^m + \sum_{i=1}^m e_i,
\]

and similarly for the lattice \( \frac{1}{k} \mathbb{Z}^m \). We index the different lattices as \( Z_I, I \in \mathbb{Z}_2^m \), ordered by \((0, \ldots, 0), (1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, 1, 1, \ldots, 1), (1, \ldots, 1)\).

Sometimes we will denote the indices by \( 1, \ldots, 2^m \) according to this order. For each \( I \in \mathbb{Z}_2^m \), denote

\[
L_I := (Z_I + [-2/k, 2/k]^m) \cap [0, 1]^m,
\]

\[
S_I := (Z_I + (-3/k, 3/k)^m) \cap [0, 1]^m.
\]

Note that \( \bigcup L_I = [0, 1]^m \) and that \( L_I \) may only intersect \( L_J \) at its boundary.

We now define diffeomorphisms \( \Phi_I(x, y) = (x + \zeta_I(x, y), y) \), such that \( \Phi = \Phi_{2^m} \circ \cdots \circ \Phi_1 \),

\[
\Phi_I \circ \cdots \circ \Phi_1 |_{[0,1) \times S_I} = \Phi_I,
\]

\[
\text{supp}(\zeta_I) \subset (0, 1) \times S_I,
\]

and

\[
0 \leq \zeta_I \leq C, \quad -1 + C^{-1} < \partial_x \zeta_I < C, \quad |\partial_y \zeta_I| < Ck, \quad \text{for some } C \text{ independent of } k. \tag{4.3}
\]

Let \( \chi_I(y) \) be a bump function such that \( \chi_I |_{L_I} = 1 \), \( \text{supp} \chi_I \subset S_I \) and \( |d\chi_I| < Ck \). Define

\[
\zeta_I(x, y) = \zeta_I(x, y) \chi_I(y).
\]

For \( I = 2, \ldots, 2^m - 1 \), define

\[
\Phi_I := \Phi \circ \Phi_I^{-1} \circ \cdots \circ \Phi_{I-1}^{-1} = (x + \zeta_I(x, y), y),
\]

and then

\[
\zeta_I(x, y) = \zeta_I(x, y) \chi_I(y).
\]

Finally, define

\[
\Phi_{2^m} := \Phi \circ \Phi_{2^m-1}^{-1} \circ \cdots \circ \Phi_{2^m-1}^{-1}.
\]

A direct calculation shows that \( \Phi_I \) satisfies (4.1)-(4.3).

In the rest of this section we are going to prove that \( \text{dist}_s(\Phi_I, \text{Id}) = o(1) \). This relies only on properties (4.2)-(4.3), hence the result also applies to \( \Phi_I \), for all \( I \in \mathbb{Z}_2^m \), since \( \zeta_I \) satisfies the same assumptions.
4.2 Step II: Squeezing the strips

**Lemma 4.2** Fix $\alpha \gg 1$. There exists a diffeomorphism $\Psi \in \text{Diff}_c(\mathbb{R}^{1+m})$, $\Psi(x, y) = (x, \psi(x, y))$, such that

$$\psi(x, y) = e^{-\alpha} (y - z) + z,$$

for every $x \in [0, 1]$ and $y \in S_1$ such that $z \in \frac{8}{k} \mathbb{Z}^m$ is the closest element to $y$ in $\frac{8}{k} \mathbb{Z}^m$. Moreover,

$$\text{dist}_s(\Psi, \text{Id}) \lesssim \alpha k^{-1-s}.$$  \hspace{1cm} (4.5)

**Proof**: Let $u_1 \in C^\infty([-4, 4]^m)$, such that $u_1(y) = -y$ for $y \in [-3, 3]^m$, and extend periodically to $\mathbb{R}^m$. Let $\chi \in C^\infty(\mathbb{R}^{1+m})$ such that $\chi \equiv 1$ on $[0, 1]^{1+m}$. Define $u_k(x, y) := \frac{k}{\alpha} u_1(ky) \chi(x, y)$. The proof continues in the same way as the proof of Lemma 3.2. \hspace{1cm} ■

Note that in $[0, 1]^{1+m}$, $\psi$ is independent of $x$. Therefore, slightly abusing notation, we write

$$\Psi(x, y) = (x, \psi(y)), \quad \Psi^{-1}(x, y) = (x, \psi^{-1}(y)).$$

We will later have $\alpha$ depend on $k$.

4.3 Step III: Affine homotopy

**Lemma 4.3**

$$\text{dist}_s(\Gamma, \text{Id}) \lesssim k^{m/2} \lambda^m/2-s = k^s e^{-(m/2-s)s}.$$

where $\Gamma = \Psi \circ \Phi_1 \circ \Psi^{-1}$.

**Proof**: Note that

$$\Gamma = (x + \zeta_1(x, \psi^{-1}(y)), y),$$

and denote

$$\xi(x, y) := \zeta_1(x, \psi^{-1}(y)), \quad \gamma(x, y) = x + \zeta_1(x, \psi^{-1}(y)).$$

It follows from the definitions of $\zeta_1$ (4.2) and $\psi$ (4.4) that $\xi$ is supported inside $(0, 1) \times \psi(S_1)$, i.e. inside $\approx k^m$ "tubes" which are translations of $(0, 1) \times [-3\lambda, 3\lambda]^m$, where $\lambda = e^{-\alpha}/k$. In particular,

$$\text{Vol}(\text{supp} \xi) \lesssim k^m \lambda^m.$$  \hspace{1cm} (4.6)

Furthermore, as in (3.8), we have from (4.3) that

$$0 \leq \xi \leq C, \quad -1 + C^{-1} < \partial_x \xi < C, \quad |\partial_y \xi| < C \lambda^{-1}.$$  \hspace{1cm} (4.7)

Consider now an affine homotopy $\Gamma_t$ from $\text{Id}$ to $\Gamma$, that is,

$$\Gamma_t(x, y) = (x + t\xi(x, y), y).$$

The same calculation as in Lemma 3.10 using the estimates (4.6)-(4.7), yields the wanted bound on $\text{dist}_s(\text{Id}, \Gamma)$. \hspace{1cm} ■

We conclude now the proof of Theorem 4.1. We showed that

$$\Phi_1 = \Psi^{-1} \circ \Gamma \circ \Psi,$$
where (following Lemmata 4.2–4.3)

\[ \text{dist}_s(\Psi, \text{Id}) \leq \alpha k^{-(1-s)}, \quad \text{dist}_s(I, \text{Id}) \leq k^\epsilon e^{-(m/2-s)\alpha}. \]

Recall that \( m = n - 1 \geq 2 \) by hypothesis. If we choose, say \( \alpha = (\log k)^2 \), we have, for any \( s < 1 \),

\[ \text{dist}_s(\Psi, \text{Id}) \leq (\log k)^2 k^{-(1-s)} = o(1), \]

\[ \text{dist}_s(I, \text{Id}) \leq k^\epsilon e^{-(m/2-s)\log k} = o(1), \]

and therefore \( \text{dist}_s(\Phi_1, \text{Id}) = o(1) \), which completes the proof.

5 The construction for \( W^{s,p}(\mathbb{R}^n) \). for \( n \geq 2 \).

In this section we explain how to modify the arguments presented above in order to extend our earlier construction to the induced \( W^{s,p} \) geodesic distance on \( \text{Diff}_c(\mathbb{R}^n) \) for \( n \geq 2 \).

**Theorem 5.1** Let \( n \geq 2 \), and denote by \((x, y)\) the coordinates on \( \mathbb{R}^n \), where \( x \in \mathbb{R} \) and \( y \in \mathbb{R}^m \) for \( m = n - 1 \). Let \( \zeta \in C^\infty_c((0, 1)^n) \) satisfying \( \zeta \geq 0, \partial_1 \zeta > -1 \). Denote \( \phi(x, y) = x + \zeta(x, y) \). Define \( \Phi \in \text{Diff}_c(\mathbb{R}^n) \) by \( \Phi(x, y) = (\phi(x, y), y) \). Then dist\(_{s,p}(\Phi, \text{Id}) = 0 \) for every \( s \in [0, 1) \) and \( p \geq 1 \) such that \( sp < n \).

As explained at the end of Section 2, this will complete the proof of Theorem 2.4.

We will use the interpolation inequality of Proposition 2.3 to estimate \( W^{s,p} \)-norms. This is not valid for \( p = 1 \), but for functions \( u \) with compact support, it follows easily from the definition (2.2) and Hölder’s inequality that \( \|u\|_{s,1} \leq C(q, \text{supp}(u))\|u\|_{s,q} \) for every \( q > 1 \), so the \( p = 1 \) case follows from estimating \( \|u\|_{s,q} \) for \( q > 1 \), for \( q \) close enough to 1 (in the construction below the various vector fields are independent of the exponent).

**Proof:**

1. **Splitting into strips and squeezing the strips**

   Fix \( k \in \mathbb{N} \). We start exactly as in Section 4.1 by writing \( \Phi = \Phi_{2^m} \circ \ldots \circ \Phi_1 \), where \( \Phi_I \) satisfies (4.2), (4.3) for \( I = 1, \ldots, 2^m \).

   It now suffices to show that dist\(_{s,p}(\text{Id}, \Phi_1) = o(1) \) as \( k \to \infty \), at a rate that depends only on the constants in (4.2), (4.3), and that thus applies to \( \Phi_2, \ldots, \Phi_{2^m} \) as well.

   To do this, we start with the the (higher-dimensional) squeezing diffeomorphism \( \Psi \) from Lemma 4.2. Then the interpolation inequality from Proposition 2.3 yields

\[ \text{dist}_{s,p}(\Psi, \text{Id}) \leq \alpha k^{-(1-s)} \quad \text{for all} \ p \in (1, \infty). \tag{5.1} \]

2. **Flowing along the squeezed strips.**

   We will now follow the procedure of Section 3 and write

\[ \Phi_1 = \Gamma^{-1} \circ \Psi^{-1} \circ \Theta \circ \Psi, \quad \Gamma, \Theta, \Psi \in \text{Diff}_c(\mathbb{R}^2), \tag{5.2} \]
where the construction of $\Theta, \Gamma$ and accompanying estimates closely follow the 2-dimensional constructions in Sections 3.3 and 3.4.

In more detail, to define $\Theta$, we first define $\tilde{\theta}(t, x, y)$ and $u(t, x, y)$ as in (3.10) and (3.11), with the only difference that now $y \in \mathbb{R}^{n-1}$. We then define $u_\delta$ as in (3.14), by convolving $u$ (in the $x$ variable only) with a mollifier $\eta_\delta$. Finally, we let $\theta(t, x, y)$ solve the ODE (3.18), and we define $\Theta(x, y) = (\theta(x, y, 1), y)$.

Then Lemma 3.3 holds as is, and in Lemma 3.4, (3.15) holds and (3.16) becomes

$$\|u_\delta\|_{W^{s, p}} \lesssim \frac{k^{n-1} \lambda^{n-s}}{\delta^{(p-1)s}}.$$ 

hence

$$\text{dist}_{s,p}(\Psi, \text{Id}) \lesssim \frac{k^{(n-1)/p} \lambda^{(n-s)/p}}{\delta^{(p-1)s/p}}.$$

3. Error correction – affine homotopy

We define $\Gamma$ by (5.2), and we estimate $\text{dist}_{s,p}(\text{Id}, \Gamma)$ by using an affine homotopy. Lemmas 3.5-3.7 hold as is, hence Corollaries 3.8-3.9 as well. Lemma 3.10 holds as well, yielding

$$\text{dist}_{s,p}(\Gamma, \text{Id}) \lesssim \frac{\delta^{1-s}}{\gamma^{1-s}} k^s.$$

The estimate is independent of $p \in (1, \infty)$ and $n$ as a consequence of the fact that the velocity field $u_t$, $0 \leq t \leq 1$ associated to the affine homotopy (which in fact does not depend on $t$) satisfies estimates that are uniform in $p$ and $n$. This follows from easy modifications of the proofs of (3.44), (3.45). The constant in the above inequality does depend on $p$ through the dependence on the constant in the interpolation inequality.

4. Conclusion of the proof

Again, choosing

$$\alpha = (\log k)^2, \quad \lambda = \frac{1}{k^{1+\log k}}, \quad \delta = \frac{1}{k^{\log k+\sqrt{\log k}},}$$

we have, for any $s < \min \{n/p, 1\}$,

$$\text{dist}_{s,p}(\Psi, \text{Id}) \lesssim (\log k)^2 k^{-(1-s)} = o(1),$$
$$\text{dist}_{s,p}(\Theta, \text{Id}) \lesssim k^{-(n/p-s)\log k + \frac{1}{p}} \sqrt{\log k^{1-s}} = o(1),$$
$$\text{dist}_{s,p}(\Gamma, \text{Id}) \lesssim k^{1-(1-s)\sqrt{\log k}} = o(1),$$

and therefore $\text{dist}_{s,p}(\Phi_1, \text{Id}) = o(1)$.

In the far subcritical regime $s < \min \{(n-1)/p, 1\}$, one can also give a simpler construction, like that of Section 4, in which the flow along the squeezed strips is carried out by an affine homotopy, and no error-correction is needed at the end.

Acknowledgements  We are grateful for Meital Kuchar for her help with the figures. This work was partially supported by the Natural Sciences and Engineering Research Council of Canada under operating grant 261955.
References

[BBHM13] M. Bauer, M. Bruveris, P. Harms, and P.W. Michor, *Geodesic distance for right invariant Sobolev metrics of fractional order on the diffeomorphism group*, Annals of Global Analysis and Geometry 44 (2013), no. 1, 5–21.

[BBM13] M. Bauer, M. Bruveris, and P.W. Michor, *Geodesic distance for right invariant Sobolev metrics of fractional order on the diffeomorphism group II*, Annals of Global Analysis and Geometry 44 (2013), no. 4, 361–368.

[BBM14] ——, *Overview of the geometries of shape spaces and diffeomorphism groups*, Journal of Mathematical Imaging and Vision 50 (2014), no. 1, 60–97.

[BM01] H. Brezis and P. Mironescu, *Gagliardo-Nirenberg, composition and products in fractional Sobolev spaces*, Journal of Evolution Equations 1 (2001), no. 4, 387–404.

[MM05] P.W. Michor and D. Mumford, *Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms*, Doc. Math. 10 (2005), 217–245.

[NPV12] E. Di Nezza, G. Palatucci, and E. Valdinoci, *Hitchhikers guide to the fractional sobolev spaces*, Bulletin des Sciences Mathématiques 136 (2012), no. 5, 521 – 573.

[Tri92] H. Triebel, *Theory of Function Spaces II*, Monographs in Mathematics, vol. 84, Birkhäuser Basel, 1992.