On Generalizations of the Nonwindowed Scattering Transform

Albert Chua[a,*], Matthew Hirn[a,b,c], Anna Little[d]

[a] *Department of Mathematics, Michigan State University, East Lansing, MI, 48824 USA*
[b] *Department of Computational Mathematics, Science & Engineering, Michigan State University, East Lansing, MI, 48824 USA*
[c] *Center for Quantum Computing, Science & Engineering Michigan State University, East Lansing, MI, 48824 USA*
[d] *Department of Mathematics and the Utah Center For Data Science, University of Utah, Salt Lake City, UT, 84112 USA*

Abstract

In this paper, we generalize finite depth wavelet scattering transforms, which we formulate as $\mathbf{L}^q(\mathbb{R}^n)$ norms of a cascade of continuous wavelet transforms (or dyadic wavelet transforms) and contractive nonlinearities. We then provide norms for these operators, prove that these operators are well-defined, and we prove are Lipschitz continuous to the action of C^2 diffeomorphisms in specific cases. Lastly, we extend our results to formulate an operator invariant to the action of rotations $R \in \mathrm{SO}(n)$ and an operator that is equivariant to the action of rotations of $R \in \mathrm{SO}(n)$.

Keywords: Wavelets, Wavelet Scattering Transform, Deformation Stability

1. Introduction

In recent years, convolutional neural networks have shown strong performance on various vision tasks like image classification [1, 2, 3, 4]. The main reason for this is that they are able to capture information at multiple scales through the use of convolutions and pooling. However, the exact method in which these networks use this information is not understood very well.

In [5], the author proposed a formulation for a simpler model for a convolutional neural network through the use of handcrafted filters, wavelets, and a series of cascading wavelet transforms. This model, called the scattering transform, and its extensions have shown success in vision tasks, quantum chemistry, manifold learning, and graph-related tasks [6, 7, 8, 9, 10].

We first provide a review of scattering transforms to motivate this paper. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a low pass filter ($\hat{\phi}(0) \neq 0$) , $\psi : \mathbb{R}^n \to \mathbb{C}$ a suitable mother wavelet ($\hat{\psi}(0) = 0$), and G^+ be a set of rotations with determinant 1. Define a set of rotations and dilations by

$$\Lambda_J := \{\lambda = 2^j r : r \in G^+, j < J\} \text{ if } J \neq \infty \tag{1}$$

and

$$\Lambda_\infty := \{2^j r : r \in G^+, j \in \mathbb{Z}\}. \tag{2}$$

For a tuple of rotations and dilations in Λ_J, define a path of length m as the tuple $p := (\lambda_1, \dots, \lambda_m)$ and let $\mathcal{P}_J$ be the set of all finite paths. The scattering propagator for $f \in \mathbf{L}^2(\mathbb{R}^n)$ and $p \in \mathcal{P}_J$ as

$$U[p]f := |||f * \psi_{\lambda_1}| * \psi_{\lambda_2}| \cdots | * \psi_{\lambda_m}|, \tag{3}$$

☆This work was partially supported by the National Science Foundation [A.C. and M.H by grant DMS-1845856; A.L. and M.H. by grant DMS-1912906], the National Institutes of Heath [M.H. by grant NIGMS-R01GM135929], and the Department of Energy [M.H. by grant DE-SC0021152].

*Corresponding author

Email addresses: chuaalbe@msu.edu (Albert Chua), mhirn@msu.edu (Matthew Hirn), little@math.utah.edu (Anna Little)

which gathers high frequency information via a cascade of wavelet transforms and nonlinearities. The scattering operator is

$$\overline{S}f(p) = \frac{1}{\mu_p} \int_{\mathbb{R}^n} U[p]f(x)\,dx \tag{4}$$

with $\mu_p := \int_{\mathbb{R}^n} U[p]\delta(x)\,dx$. Additionally, to aggregate features similar to pooling, the author of [5] define scattering operator for $f \in \mathbf{L}^2(\mathbb{R}^n)$ and $p \in \mathcal{P}_J$ as

$$S_J[p]f(x) = U[p] * \phi_{2^J}(x) = \int_{\mathbb{R}^n} U[p]f(u)\phi_{2^J}(x-u)\,du. \tag{5}$$

Additionally, the windowed scattering transform is the set of functions

$$S_J[\mathcal{P}_J]f = \{S_J[p]f\}_{p\in\mathcal{P}_J}. \tag{6}$$

This operator is similar to a convolution neural network because along each path (analogous to each layer of a convolutional neural network) a convolution, a nonlinearity is applied, and feature aggregation occurs via the low pass filter. The scattering norm for any set of paths Ω as

$$\|S_J[\Omega]f\|^2 = \sum_{p\in\Omega} \|S_J[p]x\|_2^2. \tag{7}$$

Under very stringent conditions on the mother wavelet, the author of [5] was able to prove an isometry property for the windowed scattering transform. However, the problem with the admissibility condition in [5] is that there are very few classes of wavelets that are admissible. The author of [5] mentions an analytic cubic spline Battle-Lemarié wavelet is admissible in one dimension, but provides no examples admissible classes of wavelets with $n > 1$. To our knowledge, there are no examples in the literature of wavelets that satisfy the admissibility condition when $n > 1$.

The windowed scattering transform has three important properties that are helpful for certain machine learning tasks. The first two are the following:

1. The windowed scattering transform is a well-defined mapping on $\mathbf{L}^2(\mathbb{R}^n)$ and nonexpansive. In particular, for all $f, h \in \mathbf{L}^2(\mathbb{R}^n)$,

$$\|S_J[P_J]f - S_J[P_J]h\| \leq \|f - h\|_2. \tag{8}$$

2. Let the translation of a function be denoted as $L_c x(u) = x(u-c)$. For certain classes of wavelets, we have

$$\lim_{J\to\infty} \|S_J[P_J]f - S_J[P_J]L_c f\| = 0 \tag{9}$$

for all $c \in \mathbb{R}^n$ and for all $f \in L^2(\mathbb{R}^n)$. One can think of this as local translation invariance.

Finally, for the last property, the following definition was used in [5] for Lipschitz continuity to the action of C^2 diffeomorphisms. Let $\mathcal{H}$ be a Hilbert space, $\tau \in C^2$, and define the operator $L_\tau f(x) = f(x - \tau(x))$. A translation invariant operator Φ is said to be Lipschitz continuous to the action of C^2 diffeomorphisms if for any compact $\Omega \subset \mathbb{R}^n$, there exists C_Ω such that for all $f \in \mathcal{H}$ supported in Ω and all $\tau \in C^2(\mathbb{R}^n)$, we have

$$\|\Phi(f) - \Phi(L_\tau f)\|_{\mathcal{H}} \leq C_\Omega \left(\|D\tau\|_\infty + \|D^2\tau\|_\infty\right) \|f\|_{\mathcal{H}}. \tag{10}$$

The idea is that the difference in norm is proportional to the size of $\|D\tau\|_\infty + \|D^2\tau\|_\infty$, which indicates how much L_τ deforms f. In particular, the authors of [5] show that (10) holds for the windowed scattering transform.

The concept of stability to diffeomorphisms has become a major point of study after the publication of [5]. Based on the definition above, there has been a lot of interest in exploring the stability of various operators related to machine learning and data science. For example, [7, 11] extend the scattering transform and stability of the scattering transform to graphs and compact Riemannian manifolds, respectively; the authors in [12] loosen the restriction on the regularity of τ. Other papers explore stability for different operators with desirable properties for machine learning [13, 14, 15, 16].

Although much work has appeared in recent years about operators similar to the scattering transform and about generalizations of the scattering transform, there are still some loose ends left in [5] that have not been explored yet. First, while the author of [5] does explore creating a norm for the nonwindowed scattering transform, he does not actually prove the norm is stable to diffeomorphisms. We consider a less stringent definition for stability to diffeomorphisms in the same spirit as the definition in [5] for this paper. Let V be a normed vector space. Then we say that a translation invariant operator Φ is said to be Lipschitz continuous to the action of C^2 diffeomorphisms if for any compact $\Omega \subset \mathbb{R}^n$, there exists C_Ω such that for all $f \in V$ supported in Ω and all $\tau \in C^2(\mathbb{R}^n)$, we have

$$\|\Phi(f) - \Phi(L_\tau f)\|_V \leq C_{\Omega,\tau} \|f\|_V, \tag{11}$$

where $C_{\Omega,\tau} \to 0$ as $\|D\tau\|_\infty + \|D^2\tau\|_\infty \to 0$. Like with equation (10), $\|\Phi(f) - \Phi(L_\tau f)\|_V$ depends on $\|D\tau\|_\infty + \|D^2\tau\|_\infty$.

Using this definition, we consider a slightly different problem than the author of [5] did for the nonwindowed scattering transform. The scattering transform introduced in [5] was a collection of $\mathbf{L}^1(\mathbb{R}^n)$ norms of various cascades of dyadic wavelet convolutions and modulus nonlinearities applied to a signal. However, the use of $\mathbf{L}^q(\mathbb{R}^n)$ with $q \in [1,2]$ have meaningful physical interpretations for quantum chemistry based on [9, 10]. Here, we extend the definition of the scattering transform to the continuous wavelet transform and for $\mathbf{L}^q(\mathbb{R}^n)$ norms. For the continuous wavelet transform, the one layer wavelet scattering transform with $\mathbf{L}^q(\mathbb{R}^n)$ norm is the function $S_{\mathrm{cont},q} : \mathbb{R}_+ \to \mathbb{R}$ defined as:

$$\forall\, \lambda \in \mathbb{R}_+, \quad S_{\mathrm{cont},q} f(\lambda) := \|f * \psi_\lambda\|_q^q. \tag{12}$$

Similarly, the one layer wavelet scattering transform for the dyadic wavelet transform is the function $S_{\mathrm{dyad},q} f : \mathbb{Z} \to \mathbb{R}$ defined as:

$$\forall\, j \in \mathbb{Z}, \quad S_{\mathrm{dyad},q} f(j) := \|f * \psi_j\|_q^q. \tag{13}$$

More generally, the m-layer wavelet scattering transforms $S^m_{\mathrm{cont},q} f : \mathbb{R}^m_+ \to \mathbb{R}$ and $S^m_{\mathrm{dyad},q} f : \mathbb{Z}^m \to \mathbb{R}$ are defined as

$$S^m_{\mathrm{cont},q} f(\lambda_1, \ldots, \lambda_m) := \| | | |f * \psi_{\lambda_1}| * \psi_{\lambda_2}| * \cdots | * \psi_{\lambda_m}\|_q^q, \tag{14}$$

$$S^m_{\mathrm{dyad},q} f(j_1, \ldots, j_m) := \| | | |f * \psi_{j_1}| * \psi_{j_2}| * \cdots | * \psi_{j_m}\|_q^q. \tag{15}$$

This is similar to working with a windowed scattering transform with a finite number of layers. However, our operator is different from the operator S_J in [5] because it does not contain the filter A_J to aggregate low frequency information, so the scale parameter in our formulation is not bounded above or below. With this formulation, we create different norms and prove Lipschitz continuity to C^2 diffeomorphisms that holds for an arbitrary, finite number of layers.

Remark 1. We can replace all the modulus operators with any contraction mapping (or use different contraction mappings in each layer) in the definition above, and all the proofs in the rest of this paper will still work. In particular, the modulus can be replaced with a complex version of the rectified linear unit (ReLU) nonlinearity, $\max(0, \mathrm{Re}(a_i))_{i=1,\ldots,n}$ for $a \in \mathbb{C}^n$, which is a popular choice for complex neural networks. Nonetheless, we will use the modulus operator throughout this paper without any loss of generality.

We provide a general roadmap for this paper. Section 2 will cover notation, basic properties about wavelets and the wavelet scattering operator, and harmonic analysis that will be necessary for the paper. In Section 3, we provide norms for an m-layer wavelet scattering transforms and prove that the operators are well defined mappings into specific spaces when $1 \leq q \leq 2$. For Section 4, we explore conditions under which the m-layer scattering transform is stable to dilations, and we generalize our results to diffeomorphisms in Section 5. Lastly, in Section 6, we formulate two new translation invariant operators that are stable to diffeomorphisms. The first is rotation equivariant, and the second is rotation invariant. Our contributions include, but are not limited to, the following:

- We formulate an extension of the dyadic wavelet scattering operator for a finite, arbitrary number of layers with parameter $q \in [1,2]$ by applying $\mathbf{L}^q(\mathbb{R}^n)$ norms to $U[p]f$ instead of $\mathbf{L}^1(\mathbb{R}^n)$ norms. Additionally, we formulate a wavelet scattering operator with $q \in [1,2]$ that uses a continuous scale parameter, like the continuous wavelet transform.
- We create a new finite depth scattering norm using dyadic and continuous scales in the case when $q \in [1,2]$, and prove that the mappings are well defined and provide theoretical justification for a broader class of wavelets that make the scattering transform Lipchitz continuous to the action of C^2 diffeomorphisms. However, the trade-off is that our stability bound depends on the number of layers.
- We provide a condition for norm equivalence in the case of $q = 2$ that is less stringent.
- In the case of $q \in (1,2]$, we prove that our norm is stable to diffeomorphisms $\tau \in C^2(\mathbb{R}^n)$ provided that $\|\tau\|_\infty < \frac{1}{2n}$ and provided that the wavelet and its first and second partial derivatives have sufficient decay. In the case of $q = 1$, we show stability to dilations.
- We extend our formulation to include invariance or equivariance to the action of rotations $R \in \mathrm{SO}(n)$.

2. Notation and Basic Properties

We start by providing basic notation that we will use in this paper and proceed to give basic definitions and properties that will be necessary for our results.

2.1. Function Spaces

Set $\mathbb{R}_+$ to be the positive real numbers, i.e.,

$$\mathbb{R}_+ := (0, \infty)\,.$$

The gradient of a function $f : \mathbb{R}^n \to \mathbb{C}$ is given by ∇f, the Jacobian of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is given by Df, and the Hessian is given by $D^2 f$.

For $1 \leq q < \infty$, the $\mathbf{L}^q(\mathbb{R}^n)$ norm of a function $f : \mathbb{R}^n \to \mathbb{C}$ is

$$\|f\|_q := \left[\int_{\mathbb{R}^n} |f(x)|^q \, dx\right]^{1/q}\,.$$

When $q = \infty$,

$$\|f\|_\infty = \text{ess sup}|f|.$$

Additionally, $\|\Delta f\|_\infty = \sup_{x,y \in \mathbb{R}^d} |f(x) - y(y)|$.

The Fourier transform of a function $f \in \mathbf{L}^1(\mathbb{R}^n)$ is the function $\widehat{f} \in \mathbf{L}^\infty(\mathbb{R}^n)$ defined as:

$$\forall\, \omega \in \mathbb{R}^n\,, \quad \widehat{f}(\omega) := \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\omega} \, dx\,.$$

The Hilbert transform of a function f is denoted by Hf and is defined as:

$$Hf(x) := \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} \frac{f(y)}{x-y} \, dy\,.$$

The map H is a convolution operator in which f is convolved against the function $1/x$. We note that

$$H : \mathbf{L}^q(\mathbb{R}) \to \mathbf{L}^q(\mathbb{R}) , \quad \forall 1 < q < \infty ,$$

however the result is not true for $q = 1$, i.e., if $f \in \mathbf{L}^1(\mathbb{R})$ it is not necessarily true that $Hf \in \mathbf{L}^1(\mathbb{R})$. We thus introduce the Hardy space. We denote the Hardy space as $\mathbf{H}^1(\mathbb{R})$ and it consists of those functions $f \in \mathbf{L}^1(\mathbb{R})$ such that $Hf \in \mathbf{L}^1(\mathbb{R})$ as well. For $f \in \mathbf{H}^1(\mathbb{R})$ the Hardy space norm is $\|f\|_{\mathbf{H}^1(\mathbb{R})}$, which we define as (see Corollary 2.4.7 of [17])

$$\|f\|_{\mathbf{H}^1(\mathbb{R})} := \|f\|_1 + \|Hf\|_1 . \tag{16}$$

One can show that if $f \in \mathbf{H}^1(\mathbb{R})$, then f must necessarily have zero average. An important property of the Hilbert transform and convolution is the following:

$$H(f * g) = Hf * g = f * Hg , \quad f \in \mathbf{L}^p(\mathbb{R}) , g \in \mathbf{L}^q(\mathbb{R}) , \quad 1 < \frac{1}{p} + \frac{1}{q} .$$

We also note that the Fourier transform of Hf is the following:

$$\widehat{Hf}(\omega) = \begin{cases} +i\widehat{f}(\omega) & \omega < 0 \\ -i\widehat{f}(\omega) & \omega > 0 \end{cases} \tag{17}$$

We have a similar definition for Hardy spaces when $n \geq 2$. For $1 \leq j \leq n$, define the j^{th} Riesz transform as

$$R_j f(x) = \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} \frac{x_j - y_j}{|x-y|^{n+1}} f(y)\, dy , \tag{18}$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. The Hardy space $f \in \mathbf{H}^1(\mathbb{R}^n)$ consists of functions f such that $f \in \mathbf{L}^1(\mathbb{R}^n)$ and $R_j f \in \mathbf{L}^1(\mathbb{R}^n)$ for $1 \leq j \leq n$ as well. For $f \in \mathbf{H}^1(\mathbb{R}^n)$ the Hardy space norm is $\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$, which we define as (see Corollary 2.4.7 of [17])

$$\|f\|_{\mathbf{H}^1(\mathbb{R}^n)} := \|f\|_1 + \sum_{j=1}^{n} \|R_j f\|_1 . \tag{19}$$

We have a similar multiplier property for the Fourier Transform of each Riesz Transform:

$$\widehat{R_j f}(\omega) = -i \frac{\omega_j}{|\omega|} \hat{f}(\omega) . \tag{20}$$

2.2. *Wavelets*

We let $\psi \in \mathbf{L}^1(\mathbb{R}^n) \cap \mathbf{L}^2(\mathbb{R}^n)$ be a wavelet, which means it is a function that is localized in both space and frequency and has zero average, i.e.,

$$\int_{\mathbb{R}^n} \psi(x)\, du = 0 .$$

For a continuous dilation parameter $\lambda \in \mathbb{R}_+$ we define the dilations of ψ as:

$$\forall \lambda \in \mathbb{R}_+ , \quad \psi_\lambda(x) := \lambda^{-n/2} \psi(\lambda^{-1} x) ,$$

which preserves the $\mathbf{L}^2(\mathbb{R}^n)$ norm of ψ:

$$\|\psi_\lambda\|_2 = \|\psi\|_2 , \quad \forall \lambda \in \mathbb{R}_+ .$$

The Fourier transform of ψ_λ is:

$$\widehat{\psi}_\lambda(\omega) = \lambda^{n/2}\widehat{\psi}(\lambda\omega)\,.$$

Assume $f \in \mathbf{L}^2(\mathbb{R}^n)$. The continuous wavelet transform $\mathcal{W}f \in \mathbf{L}^2(\mathbb{R}^n \times \mathbb{R}_+)$ is defined as:

$$\forall\,(x,\lambda) \in \mathbb{R}^n \times \mathbb{R}_+\,, \quad \mathcal{W}f(x,\lambda) := f * \psi_\lambda(x)\,.$$

Furthermore, assume that ψ satisfies the following admissibility condition

$$\int_0^\infty \frac{|\widehat{\psi}(\lambda\omega)|^2}{\lambda}\,d\lambda = \mathcal{C}_\psi\,, \quad \forall\,\omega \in \mathbb{R}^n \setminus \{0\}\,, \tag{21}$$

for some $\mathcal{C}_\psi > 0$. Then we will say that ψ is a Littlewood-Paley wavelet for the continuous wavelet transform. If ψ satisfies (21), one can show that the norm $\mathcal{W}f$ computed with a weighted measure $(dx, d\lambda/\lambda^{n+1})$ on $\mathbb{R}^n \times \mathbb{R}_+$ is well defined:

$$\|\mathcal{W}f\|^2_{\mathbf{L}^2(\mathbb{R}^n\times\mathbb{R}_+)} := \int_0^\infty \int_{\mathbb{R}^n} |\mathcal{W}f(x,\lambda)|^2\,dx\,\frac{d\lambda}{\lambda^{n+1}} = \int_0^\infty \int_{\mathbb{R}^n} |f * \psi_\lambda(x)|^2\,dx\,\frac{d\lambda}{\lambda^{n+1}} = \int_0^\infty \|f*\psi_\lambda\|_2^2\,\frac{d\lambda}{\lambda^{n+1}}\,.$$

We note, in fact, that one can show:

$$\|\mathcal{W}f\|^2_{\mathbf{L}^2(\mathbb{R}^n\times\mathbb{R}_+)} = \beta\cdot\mathcal{C}_\psi\|f\|_2^2\,.$$

where

$$\beta = \begin{cases} 1 & \text{if } \psi \text{ is real valued} \\ 1/2 & \text{if } \psi \text{ is complex analytic} \end{cases}\,. \tag{22}$$

For a dyadic dilation parameter $j \in \mathbb{Z}$ we define dilations of ψ as:

$$\forall\, j \in \mathbb{Z}\,, \quad \psi_j(x) = 2^{-nj}\psi(2^{-j}x)\,,$$

which in this case preserves the $\mathbf{L}^1(\mathbb{R}^n)$ norm of ψ:

$$\|\psi_j\|_1 = \|\psi\|_1\,, \quad \forall\, j \in \mathbb{Z}\,.$$

The Fourier transform of ψ_j is:

$$\widehat{\psi}_j(\omega) = \widehat{\psi}(2^j\omega)\,.$$

For a function $f \in \mathbf{L}^2(\mathbb{R}^n)$ we define the dyadic wavelet transform $Wf \in \ell^2(\mathbf{L}^2(\mathbb{R}^n))$ as

$$Wf = \left(f * \psi_j\right)_{j\in\mathbb{Z}}\,.$$

Assume that ψ satisfies

$$\sum_{j\in\mathbb{Z}} |\widehat{\psi}(2^j\omega)|^2 = \hat{C}_\psi, \quad \forall\omega \in \mathbb{R}^n \setminus \{0\}\,, \tag{23}$$

for some $\hat{C}_\psi > 0$. Then we will say that ψ is a Littlewood-Paley wavelet for the dyadic wavelet transform. If ψ satisfies (23), one can show that the norm Wf given below is well defined:

$$\|Wf\|^2_{\ell^2(\mathbf{L}^2(\mathbb{R}))} := \sum_{j\in\mathbb{Z}} \|f*\psi_j\|_2^2\,.$$

In fact, we have the following norm equivalence:

$$\|Wf\|^2_{\ell^2(\mathbf{L}^2(\mathbb{R}))} = \beta\cdot\hat{C}_\psi\|f\|_2^2\,,$$

where β is defined in (22).

We also remark that the Hilbert transform commutes with dilations, so in particular:

$$H(\psi_\lambda) = H(\psi)_\lambda \quad \text{and} \quad H(\psi_j) = H(\psi)_j\,.$$

Additionally, using the calculation of $\widehat{Hf}$ in (17) we see that

$$H\psi = -i\psi\,, \quad \text{if } \psi \text{ is complex analytic.}$$

2.3. Operator Valued Spaces

Consider a Banach space $\mathcal{B}$. Suppose $f : \mathbb{R}^n \to \mathcal{B}$ and $x \to \|f(x)\|_{\mathcal{B}}$ is measurable in the Lebesgue sense. Define $\mathbf{L}^p_{\mathcal{B}}(\mathbb{R}^n)$ for $1 \le p < \infty$ to be

$$\|f\|^p_{\mathbf{L}^p_{\mathcal{B}}(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \|f(x)\|^p_{\mathcal{B}}\, dx\,.$$

Also, for $1 \le p < \infty$, define

$$\|f\|_{\mathbf{L}^{p,\infty}_{\mathcal{B}}(\mathbb{R}^n)} = \sup_{\delta>0} \delta \cdot m(\{x \in \mathbb{R}^n : \|f(x)\|_{\mathcal{B}} > \delta\})^{1/p}\,.$$

We also have the following relation:

$$\|f\|_{\mathbf{L}^{p,\infty}_{\mathcal{B}}(\mathbb{R}^n)} \le \|f\|_{\mathbf{L}^p_{\mathcal{B}}(\mathbb{R}^n)}\,.$$

Note that for $f : \mathbb{R}^n \to \mathbb{R}^n$,

$$\|f\|^p_{\mathbf{L}^p_{\mathbb{R}^n}(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \|f(x)\|^p_{\mathbb{R}^n}\, dx = \int_{\mathbb{R}^n} |f(x)|^p\, dx = \|f\|^p_p\,.$$

3. Wavelet Scattering is a Bounded Operator

In this section we explore for which $q > 0$ and $m \ge 1$ the wavelet scattering transforms $S^m_{\mathrm{cont},q} f$ and $S^m_{\mathrm{dyad},q} f$ are well-defined as functions in some Banach space (i.e., have finite norm), and under what circumstances.

Let ψ be a wavelet and define $\psi_t(x) = t^{-n}\psi(x/t)$ for $t \in (0,\infty)$. We assume that ψ has the following properties:

$$|\psi(x)| \le A(1+|x|)^{-n-\varepsilon}\,, \tag{24}$$

$$\int_{\mathbb{R}^n} |\psi(x+h) - \psi(x)|\, dx \le A|h|^{\varepsilon'}\,, \tag{25}$$

for some constants $A, \varepsilon', \varepsilon > 0$ and for all $h \neq 0$.

Consider the Littlewood-Paley G-function

$$G_\psi(f)(x) = \left(\int_{(0,\infty)} |f * \psi_t(x)|^2 \frac{dt}{t} \right)^{1/2}\,. \tag{26}$$

Let $\mathcal{B} = \mathbf{L}^2\left((0,\infty), \frac{dt}{t}\right)$. We can rewrite this as a Bochner integral by considering the function $K(x) = (\psi_t(x))_{t>0}$. This is a mapping $K : \mathbb{R}^n \to \mathcal{B}$ and the function $x \to \|K(x)\|_{\mathcal{B}}$ is measurable. Also, if we let

$$T(f)(x) = \left(\int_{\mathbb{R}^n} \psi_t(x-y) f(y)\, dy \right)_{t>0} = ((\psi_t * f)(x))_{t>0}\,,$$

we observe that

$$G_\psi(f)(x) = \|T(f)(x)\|_{\mathcal{B}}$$

and

$$\|G_\psi(f)\|^p_p = \|T(f)\|^p_{L^p_{\mathcal{B}}(\mathbb{R}^n)}\,.$$

The two properties above for the wavelet ψ imply that

$$\|K(x)\|_{\mathcal{B}} \le \frac{c_n A}{|x|^n}\,, \tag{27}$$

and

$$\sup_{h \in \mathbb{R}^n \setminus \{0\}} \int_{\mathbb{R}^n} \|K(x+h) - K(x)\|_{\mathcal{B}} dx \le c'_n A\,, \tag{28}$$

where c_n and c'_n depend only on n.

Remark 2. For the rest of this paper, we will write G in place of G_ψ when referring to the G-function because the dependence on the mother wavelet is clear.

We have the following result taken from Problem 6.1.4 of [18] and from Chapter V of [19].

Lemma 1 ([18, 19])**.** *Assume that ψ is defined as above and satisfies (27) and (28). Then the operator G is bounded from $\mathbf{L}^2(\mathbb{R}^n)$ to $\mathbf{L}^2(\mathbb{R}^n)$. Also, for $p \in (1, \infty)$ and $\mathcal{B} = \mathbf{L}^2(\mathbb{R}_+, dt/t)$, we have*

$$\|Tf\|_{\mathbf{L}^p_{\mathcal{B}}(\mathbb{R}^n)} \leq C_n(A+B)\max(p,(p-1)^{-1})\|f\|_{\mathbf{L}^p(\mathbb{R}^n)}\,, \tag{29}$$

where $B = \|G\|_{\mathbf{L}^2(\mathbb{R}^n)\to\mathbf{L}^2(\mathbb{R}^n)}$ and C_n only depends on n. For all $f \in \mathbf{L}^1(\mathbb{R})$, we also have

$$\|Tf\|_{\mathbf{L}^{1,\infty}_{\mathcal{B}}(\mathbb{R}^n)} \leq C'_n(A+B)\|f\|_{\mathbf{L}^1(\mathbb{R}^n)} \tag{30}$$

and

$$\|Tf\|_{\mathbf{L}^1_{\mathcal{B}}(\mathbb{R}^n)} \leq C'_n(A+B)\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}\,, \tag{31}$$

where C'_n only depends on n.

Remark 3. We can also formulate similar bounds for the Littlewod-Paley $\mathfrak{g}$ operator

$$\mathfrak{g}(f)(x) := \left[\sum_{j\in\mathbb{Z}} |\psi_j * f(x)|^2\right]^{1/2} \tag{32}$$

using similar arguments.

Remark 4. Let ψ be a wavelet that has properties (24) and (25). Then with the $\mathbf{L}^2$ normalized dilations, the Littlewood-Paley G-function can be written as:

$$G(f)(x) = \left[\int_0^\infty |f * \psi_\lambda(x)|^2 \frac{d\lambda}{\lambda^{n+1}}\right]^{1/2}. \tag{33}$$

Note that the λ measure for $G(f)$ matches the measure in defining the norm of $W_{\mathrm{cont}}f$.

3.1. *The $\mathbf{L}^2(\mathbb{R}^n)$ Wavelet Scattering Transform*

In this section we prove the $\mathbf{L}^2(\mathbb{R}^n)$ scattering transforms are bounded operators. More specifically, we prove that $S^m_{\mathrm{cont},2} : \mathbf{L}^2(\mathbb{R}^n) \to \mathbf{L}^1(\mathbb{R}^m_+)$, where $\mathbf{L}^1(\mathbb{R}^m_+)$ has the weighted measure defined by

$$\|S^m_{\mathrm{cont},2}f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} := \int_0^\infty \cdots \int_0^\infty |S^m_{\mathrm{cont},2}f(\lambda_1,\ldots,\lambda_m)|\,\frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_m}{\lambda_m^{n+1}}$$

and we show that $\|S^m_{\mathrm{cont},2}f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \leq C\|f\|_{\mathbf{L}^2(\mathbb{R}^n)}$. We also show that $S^m_{\mathrm{dyad},2} : \mathbf{L}^2(\mathbb{R}^n) \to \ell^1(\mathbb{Z}^m)$, where

$$\|S^m_{\mathrm{dyad},2}f\|_{\ell^1(\mathbb{Z}^m)} := \sum_{j_m\in\mathbb{Z}} \cdots \sum_{j_1\in\mathbb{Z}} |S^m_{\mathrm{dyad},2}f(j_1,\ldots,j_m)|^2.$$

Proposition 2. *For any wavelet satisfying* (24) *and* (25)*, we have $S^m_{cont,2} : \mathbf{L}^2(\mathbb{R}^n) \to \mathbf{L}^1(\mathbb{R}^m_+)$ and $S^m_{dyad,2} : \mathbf{L}^2(\mathbb{R}^n) \to \ell^1(\mathbb{Z}^m)$.*

Proof. The proof of the dyadic case is essentially identical to the proof given below and is thus omitted. The case of $m=1$ follows by an application of Fubini's Theorem:

$$
\begin{aligned}
\|S_{\text{cont},2}f\|_{\mathbf{L}^1(\mathbb{R}_+)} &= \int_0^\infty \|f*\psi_\lambda\|_2^2 \frac{d\lambda}{\lambda^{n+1}} \\
&= \int_0^\infty \int_{\mathbb{R}^n} |(f*\psi_\lambda)(x)|^2\, dx \frac{d\lambda}{\lambda^{n+1}} \\
&= \int_{\mathbb{R}^n} |G(f)(x)|^2\, dx \\
&\leq C\|f\|_2^2
\end{aligned}
$$

by boundedness of the G-function. Now we proceed by using induction. Assume that we have $\|S^m_{\text{cont},2}f\|^2_{\mathbf{L}^1(\mathbb{R}^m_+)} \leq C_m\|f\|_2^2$. Let $\mathcal{W}_t f = f*\psi_t$, define $Mf = |f|$, and $U_\lambda = M\mathcal{W}_\lambda$ for notational brevity. Then notice that

$$
\||||f*\psi_{\lambda_1}|*\psi_{\lambda_2}|*\cdots*\psi_{\lambda_m}|*\psi_{\lambda_{m+1}}\|_2^2 = \|\mathcal{W}_{\lambda_{m+1}}U_{\lambda_m}\cdots U_{\lambda_1}f\|_2^2.
$$

$$
\begin{aligned}
\|S^{m+1}_{\text{cont},2}f\|_{\mathbf{L}^1(\mathbb{R}^{m+1}_+)} &= \int_0^\infty\cdots\int_0^\infty \|\mathcal{W}_{\lambda_{m+1}}U_{\lambda_m}\cdots U_{\lambda_1}f\|_2^2 \frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_{m+1}}{\lambda_{m+1}^{n+1}} \\
&= \int_0^\infty\cdots\int_0^\infty\int_0^\infty \|(U_{\lambda_m}\cdots U_{\lambda_1}f)*\psi_{\lambda_{m+1}}\|_2^2 \frac{d\lambda_{m+1}}{\lambda_{m+1}^{n+1}}\frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_m}{\lambda_m^{n+1}} \\
&= \int_0^\infty\cdots\int_0^\infty \|(U_{\lambda_m}\cdots U_{\lambda_1}f)*\psi_{\lambda_{m+1}}\|_{\mathbf{L}^1(\mathbb{R}_+)} \frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_m}{\lambda_m^{n+1}} \\
&\leq C\int_0^\infty\cdots\int_0^\infty \|U_{\lambda_m}\cdots U_{\lambda_1}f\|_2^2 \frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_m}{\lambda_m^{n+1}} \\
&= C\int_0^\infty\cdots\int_0^\infty S^m_{\text{dyad},2}(\lambda_1,\ldots,\lambda_m) \frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_m}{\lambda_m^{n+1}} \\
&\leq C^{m+1}\|f\|_2^2,
\end{aligned}
$$

where we used the induction hypothesis in the last line. This completes the proof. □

Proposition 3. *Suppose ψ is a Littlewood-Paley wavelet satisfying (24) and (25). Then $S^m_{cont,2}: \mathbf{L}^2(\mathbb{R}^n)\to\mathbf{L}^1(\mathbb{R}^m_+)$ and specifically $\|S^m_{cont,2}f\|_1 = C^m_\psi\|f\|_2^2$. Also, $S^m_{dyad,2}: \mathbf{L}^2(\mathbb{R}^n)\to\ell^1(\mathbb{Z}^m)$ and $\|S^m_{dyad,2}f\|_1 = \hat{C}^m_\psi\|f\|_2^2$.*

Proof. We only provide the proof of the continuous case again. First consider the case $m=1$. We have:

$$
\begin{aligned}
\|S_{\text{cont},2}f\|_{\mathbf{L}^1(\mathbb{R}_+)} &= \int_0^\infty \|f*\psi_\lambda\|_2^2\frac{d\lambda}{\lambda^{n+1}} \\
&= \frac{1}{(2\pi)^n}\int_0^\infty \|\hat{f}\cdot\hat{\psi}_\lambda\|_2^2\frac{d\lambda}{\lambda^{n+1}} \\
&= \frac{1}{(2\pi)^n}\int_0^\infty \left(\int_{\mathbb{R}^n}|\hat{f}(\omega)|^2|\hat{\psi}_\lambda(\omega)|^2\, d\omega\right)\frac{d\lambda}{\lambda^{n+1}} \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left(\int_0^\infty |\hat{\psi}(\lambda\omega)|^2\frac{d\lambda}{\lambda}\right)|\hat{f}(\omega)|^2\, d\omega \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left(\hat{C}_\psi|\hat{f}(\omega)|^2\right)\, d\omega \\
&= \frac{1}{(2\pi)^n}\hat{C}_\psi\|\hat{f}\|_2^2 \\
&= C_\psi\|f\|_2^2.
\end{aligned}
$$

Thus the claim holds for $m = 1$. Now assume that it holds through m. Then by the inductive hypothesis,

$$\|S^m_{\text{cont},2} f\|_{\mathbf{L}^1(\mathbb{R}_+)} = \int_0^\infty \cdots \int_0^\infty \||| f * \psi_{\lambda_1}| * \psi_{\lambda_2}| * \cdots * \psi_{\lambda_m}\|_2^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}} = C_\psi^m \|f\|_2^2.$$

Now consider the case of $m+1$. Similar to the previous proposition, we have

$$\begin{aligned}
\|S^{m+1}_{\text{cont},2} f\|_{\mathbf{L}^1(\mathbb{R}_+)} &= \int_0^\infty \cdots \int_0^\infty \left(\int_0^\infty \|(U_{\lambda_m} \cdots U_{\lambda_1} f) * \psi_{\lambda_{m+1}}\|_2^2 \frac{d\lambda_{m+1}}{\lambda_{m+1}^{n+1}} \right) \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}} \\
&= C_\psi \int_0^\infty \cdots \int_0^\infty S_{\text{cont},2} f(\lambda_1, \ldots, \lambda_m) \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}} \\
&= C_\psi \|S^m_{\text{cont},2} f\|_{\mathbf{L}^1(\mathbb{R}_+)} \\
&= C_\psi^{m+1} \|f\|_2^2.
\end{aligned}$$

Thus, the claim is proven by induction. □

3.2. *The $\mathbf{L}^1(\mathbb{R}^n)$ Wavelet Scattering Transform*

We now consider the case of $q = 1$.

3.2.1. *The One-layer $\mathbf{L}^1(\mathbb{R}^n)$ Wavelet Scattering Transform*

In this section we prove that $\mathbb{S}_{\text{cont},1} : \mathbf{H}^1(\mathbb{R}^n) \to \mathbf{L}^2(\mathbb{R}_+)$, where $\mathbf{L}^2(\mathbb{R}_+)$ has the weighted measure. In particular, we define:

$$\|S_{\text{cont},1} f\|^2_{\mathbf{L}^2(\mathbb{R}_+)} := \int_0^\infty |S_{\text{cont},1} f(\lambda)|^2 \frac{d\lambda}{\lambda^{n+1}} = \int_0^\infty \|f * \psi_\lambda\|_1^2 \frac{d\lambda}{\lambda^{n+1}},$$

and show that $\|S_{\text{cont},1} f\|_{\mathbf{L}^2(\mathbb{R}_+)} \leq C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$. We also show that $S_{\text{dyad},1} : \mathbf{H}^1(\mathbb{R}^n) \to \ell^2(\mathbb{Z})$, where

$$\|S_{\text{dyad},1} f\|^2_{\ell^2(\mathbb{Z})} := \sum_{j \in \mathbb{Z}} |S_{\text{dyad},1} f(j)|^2 = \sum_{j \in \mathbb{Z}} \|f * \psi_j\|_1^2.$$

Proposition 4. *Let ψ be a wavelet that satisfies properties (24) and (25) and let $S_{cont,1}$ and $S_{dyad,1}$ be defined as above. Then $S_{cont,1} : \mathbf{H}^1(\mathbb{R}^n) \to \mathbf{L}^2(\mathbb{R}_+)$, that is there exists a universal constant $C > 0$ such that $\|S_{cont,1} f\|_{\mathbf{L}^2(\mathbb{R}_+)} \leq C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$ for all $f \in \mathbf{H}^1(\mathbb{R}^n)$, and furthermore, $S_{dyad,1} : \mathbf{H}^1(\mathbb{R}^n) \to \ell^2(\mathbb{Z})$ with $\|S_{dyad,1} f\|_{\ell^2(\mathbb{Z})} \leq C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$.*

Proof. Let $f \in \mathbf{H}^1(\mathbb{R}^n)$ throughout the proof. For the continuous wavelet transform we have:

$$\begin{aligned}
\|S_{\text{cont},1} f\|_{\mathbf{L}^2(\mathbb{R}_+)} &= \left(\int_0^\infty \|f * \psi_\lambda\|_1^2 \frac{d\lambda}{\lambda^{n+1}} \right)^{1/2} \\
&= \left(\int_0^\infty \left(\int_{\mathbb{R}^n} |f * \psi_\lambda(x)| \, dx \right)^2 \frac{d\lambda}{\lambda^{n+1}} \right)^{1/2} \\
&\leq \left(\int_{\mathbb{R}^n} \left(\int_0^\infty |f * \psi_\lambda(x)|^2 \frac{d\lambda}{\lambda^{n+1}} \right)^{1/2} dx \right) \\
&= \int_{\mathbb{R}^n} G(f)(x) \, dx \\
&= \|G(f)\|_1 \\
&\leq C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)},
\end{aligned}$$

where in the last inequality we used Theorem 1. A similar proof follows for the dyadic case. □

3.2.2. *The two-layer $\mathbf{L}^1(\mathbb{R}^n)$ Wavelet Scattering Transform*

Now we prove that $S^2_{\text{cont},1} : \mathbf{H}^1(\mathbb{R}^n) \to \mathbf{L}^2(\mathbb{R}^2_+)$. The norm for $S^2_{\text{cont},1} f$ is:

$$\|S^2_{\text{cont},1} f\|^2_{\mathbf{L}^2(\mathbb{R}^2_+)} := \int_0^\infty \int_0^\infty |S^2_{\text{cont},1} f(\lambda_1, \lambda_2)|^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \frac{d\lambda_2}{\lambda_2^{n+1}} = \int_0^\infty \int_0^\infty \| |f * \psi_{\lambda_1}| * \psi_{\lambda_2} \|_1^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \frac{d\lambda_2}{\lambda_2^{n+1}}.$$

Instead of skipping directly to a proof for m layers, a slight technicality arises because of the Riesz Transforms. Thus, we need to add some restrictions on the choice of wavelet.

Proposition 5. *Let ψ be a wavelet with $n+3$ vanishing moments that satisfies properties (24) and (25) and let $S^2_{cont,1}$ be defined as above. Then $S^2_{cont,1} : \mathbf{H}^1(\mathbb{R}^n) \to \mathbf{L}^2(\mathbb{R}^2_+)$; that is, there exists a universal constant $C > 0$ such that $\|S^2_{cont,1} f\|_{\mathbf{L}^2(\mathbb{R}^2_+)} \le C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$ for all $f \in \mathbf{H}^1(\mathbb{R}^n)$.*

The following lemma will be useful in proving Proposition 5.

Lemma 6. *Let ψ be a wavelet that satisfies properties (24) and (25) and let $S^2_{cont,1}$ be defined as above. Then for $f \in \mathbf{H}^1(\mathbb{R}^n)$,*

$$\|S^2_{cont,1} f\|^2_{\mathbf{L}^2(\mathbb{R}^2_+)} \le C \int_0^\infty \|f * \psi_\lambda\|^2_{\mathbf{H}^1(\mathbb{R}^n)} \frac{d\lambda}{\lambda^{n+1}}.$$

An identical result follows for the dyadic case.

Proof. We'll write the norm $\|S^2_{\text{cont},1} f\|_{\mathbf{L}^2(\mathbb{R}^2_+)}$ as integrals, swap the order, and then use Theorem 1 to obtain the result:

$$\begin{aligned}
\|S^2_{\text{cont},1} f\|_{\mathbf{L}^2(\mathbb{R}_+)} &= \left(\int_0^\infty \int_0^\infty \| |f * \psi_{\lambda_1}| * \psi_{\lambda_2} \|_1^2 \frac{d\lambda_2}{\lambda_2^{n+1}} \frac{d\lambda_1}{\lambda_1^{n+1}} \right)^{1/2} \\
&= \left(\int_0^\infty \left[\int_0^\infty \left(\int_{\mathbb{R}^n} \left| |f * \psi_{\lambda_1}| * \psi_{\lambda_2}(x) \right| \, dx \right)^2 \frac{d\lambda_2}{\lambda_2^{n+1}} \right]^{2/2} \frac{d\lambda_1}{\lambda_1^{n+1}} \right)^{1/2} \\
&\le \left(\int_0^\infty \left[\int_0^\infty \left(\int_{\mathbb{R}^n} \left| |f * \psi_{\lambda_1}| * \psi_{\lambda_2}(x) \right|^2 \frac{d\lambda_2}{\lambda_2^{n+1}} \right)^{1/2} dx \right]^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \right)^{1/2} \\
&= \left(\int_0^\infty \left[\int_{\mathbb{R}^n} G_{\psi_{\lambda_1}}(|f * \psi_{\lambda_1}|)(x) \, dx \right]^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \right)^{1/2} \\
&\le \left(\int_0^\infty \|f * \psi_{\lambda_1}\|^2_{\mathbf{H}^1(\mathbb{R}^n)} \frac{d\lambda_1}{\lambda_1^{n+1}} \right)^{1/2}.
\end{aligned}$$

□

Proof of Proposition 5. The proof will be done for $n \ge 2$. Otherwise, we will replace the sum of the Reisz Transform components with a Hilbert Transform in every step of the argument. From Lemma 6 it is sufficient to bound $\left(\int_0^\infty \|f * \psi_\lambda\|^2_{\mathbf{H}^1(\mathbb{R}^n)} \lambda^{-n-1} d\lambda \right)^{1/2}$. Using the definition of the Hardy space norm in (18) plus

Proposition 4, we obtain:

$$\begin{aligned}
\left(\int_0^\infty \|f*\psi_\lambda\|^2_{\mathbf{H}^1(\mathbb{R}^n)}\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2} &= \left(\int_0^\infty\left(\|f*\psi_\lambda\|_1+\sum_{j=1}^n\|R_j(f*\psi_\lambda)\|_1\right)^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}\\
&\leq 2\left(\int_0^\infty\|f*\psi_\lambda\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}+2\left(\int_0^\infty\sum_{j=1}^n\|R_j(f*\psi_\lambda)\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}\\
&= 2\left(\int_0^\infty\|f*\psi_\lambda\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}+2\left(\sum_{j=1}^n\int_0^\infty\|R_j(f*\psi_\lambda)\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}
\end{aligned}$$

Thus the crux of the matter is to bound $\left(\int_0^\infty\|R_j(f*\psi_\lambda)\|_1^2\,\lambda^{-n-1}d\lambda\right)^{1/2}$. Notice that

$$\begin{aligned}
\left(\int_0^\infty\|R_j(f*\psi_\lambda)\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2} &= \left(\int_0^\infty\|f*R_j\psi_\lambda\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}\\
&= \left(\int_0^\infty\left[\int_{\mathbb{R}^n}|f*R_j\psi_\lambda|\,dx\right]^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}\\
&\leq \left(\int_{\mathbb{R}^n}\left[\int_0^\infty|f*R_j\psi_\lambda|^2\frac{d\lambda}{\lambda^{n+1}}\right]^{1/2}dx\right)\\
&= \int_{\mathbb{R}^n}G_{R_j\psi}(f)(x)\,dx.
\end{aligned}$$

It suffices to now show that $R_j\psi_\lambda$ is a wavelet. Suppose that ψ has $n+3$ vanishing moments in the k^{th} component. Then

$$\int_{\mathbb{R}} t_k^j\psi(t_1,t_2,\ldots,t_n)\,dt_k$$

for $j=0,\ldots,n+2$. Using properties of the Fourier Transform, we see that

$$\hat{\psi}(\omega)=(i\omega_k)^{n+3}\hat{\theta}(\omega).$$

with $\hat{\theta}\in\mathbf{L}^\infty(\mathbb{R}^n)$. In this case,

$$\lim_{|\omega|\to 0}\widehat{R_j\psi}(\omega)=\lim_{|\omega|\to 0}-i\frac{w_j^{n+3}}{|w|}\hat{\theta}(\omega)=0.$$

Thus, $R_j\psi$ is a wavelet with $O((1+|x|)^{-n-1})$ decay, and all its first and second order partial derivatives also have the same order of decay.

Picking up from before, we may once again apply Proposition 4 to obtain:

$$\int_{\mathbb{R}^n}G(|f*R_j\psi_\lambda|)(x)\,dx=\|G_{R_j\psi}(f)\|_{\mathbf{L}^1(\mathbb{R}^n)}\leq C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}.$$

Going back, we see that

$$\begin{aligned}
\left(\int_0^\infty \|f*\psi_\lambda\|^2_{\mathbf{H}^1(\mathbb{R}^n)}\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2} &\leq 2\left(\int_0^\infty\|f*\psi_\lambda\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}+2\left(\sum_{j=1}^n\int_0^\infty\|R_j(f*\psi_\lambda)\|_1^2\frac{d\lambda}{\lambda^{n+1}}\right)^{1/2}\\
&\leq 2C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}+2\sqrt{n}C\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}.
\end{aligned}$$

The proof is completed. $\square$

3.2.3. *The m-layer $\mathbf{L}^1(\mathbb{R}^n)$ Wavelet Scattering Transform*

Let $\mathcal{W}_t f = f * \psi_t$, $Mf = |f|$, and $U_t = M\mathcal{W}_t$ again. Then Now we try to prove that for $m \in \mathbb{N}$, $S^m_{\text{cont},1} : \mathbf{H}^1(\mathbb{R}^n) \to \mathbf{L}^2(\mathbb{R}^m_+)$. The norm for $S^m_{\text{cont},1} f$ is:

$$\|S^m_{\text{cont},1} f\|^2_{\mathbf{L}^2(\mathbb{R}^m_+)} := \int_0^\infty \int_0^\infty \cdots \int_0^\infty |S^m_{\text{cont},1} f(\lambda_1, \lambda_2, \ldots, \lambda_m)|^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \frac{d\lambda_2}{\lambda_2^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}}$$
$$= \int_0^\infty \int_0^\infty \cdots \int_0^\infty \|(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f) * \psi_{\lambda_m}\|_1^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \frac{d\lambda_2}{\lambda_2^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}}.$$

Theorem 7. *Let ψ be a wavelet with $n+3$ vanishing moments that satisfies properties (24) and (25) and let $\S^m_{cont,1}$ be defined as above. Then for $f \in \mathbf{H}^1(\mathbb{R})$, there exists a constant C_m such that*

$$\|S^m_{cont,1} f\|^2_{\mathbf{L}^2(\mathbb{R}^m_+)} \le C_m \|f\|^2_{\mathbf{H}^1(\mathbb{R}^n)}.$$

Additionally,

$$\|S^m_{dyad,1} f\|_{\ell^2(\mathbb{Z}^m)} \le C_m \|f\|_{\mathbf{H}^1(\mathbb{R}^n)}.$$

Proof. We proceed by induction. Propositions 4 and 5 already proved the base case is satisfied (in fact two base cases). Now, let us assume that

$$\|S^{m-1}_{\text{cont},1} f\|^2_{\mathbf{L}^2(\mathbb{R}^{m-1}_+)} \le C_{m-1} \|f\|^2_{\mathbf{H}^1(\mathbb{R}^n)}.$$

By a calculation similar to the two-layer calculation, we have

$$\begin{aligned}
\|S^m_{\text{cont},1} f\|_{\mathbf{L}^2(\mathbb{R}^m_+)} &= \left(\int_0^\infty \cdots \int_0^\infty \|(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f) * \psi_{\lambda_m}\|_1^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}} \right)^{1/2} \\
&= \left(\int_0^\infty \cdots \int_0^\infty \left(\int_{\mathbb{R}} |(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f) * \psi_{\lambda_m}| \, dx \right)^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}} \right)^{1/2} \\
&\le \left(\int_0^\infty \cdots \int_0^\infty \left(\int_{\mathbb{R}} \left[\int_0^\infty |(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f) * \psi_{\lambda_m}|^2 \frac{d\lambda_m}{\lambda_m^2} \right]^{1/2} dx \right)^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}} \right)^{1/2} \\
&= \left(\int_0^\infty \cdots \int_0^\infty \left[\int_{\mathbb{R}} G_{\lambda_m}(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f)(x) \, dx \right]^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}} \right)^{1/2} \\
&= \left(\int_0^\infty \cdots \int_0^\infty \|G(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f)\|_1^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}} \right)^{1/2} \\
&= \left(\int_0^\infty \cdots \int_0^\infty \|G(\mathcal{W}_{\lambda_{m-1}} U_{\lambda_{m-2}} \cdots U_{\lambda_1} f)\|_1^2 \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}} \right)^{1/2}
\end{aligned}$$

since the G function has a modulus already.

It follows that

$$\|S^m_{\text{cont},1} f\|_{\mathbf{L}^2(\mathbb{R}^m_+)} \le C^{1/2} \left(\int_0^\infty \cdots \int_0^\infty \|\mathcal{W}_{\lambda_{m-1}} U_{\lambda_{m-2}} \cdots U_{\lambda_1} f\|^2_{\mathbf{H}^1(\mathbb{R}^n)} \frac{d\lambda_1}{\lambda_1^{n+1}} \cdots \frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}} \right)^{1/2}.$$

Now use the definition of the $\mathbf{H}^1(\mathbb{R}^n)$ norm to write

$$\|\mathcal{W}_{\lambda_{m-1}} U_{\lambda_{m-2}} \cdots U_{\lambda_1} f\|_{\mathbf{H}^1(\mathbb{R}^n)} = \|\mathcal{W}_{\lambda_{m-1}} U_{\lambda_{m-2}} \cdots U_{\lambda_1} f\|_{\mathbf{L}^1(\mathbb{R}^n)} + \sum_{j=1}^n \left\| (R_j \mathcal{W}_{\lambda_{m-1}}) (U_{\lambda_{m-2}} \cdots U_{\lambda_1} f) \right\|_{\mathbf{L}^1(\mathbb{R}^n)}.$$

Thus, since $R_j \mathcal{W}_{\lambda_{m-1}} h = h * (R_j \psi_{\lambda_{m-1}})$, we can use our induction hypothesis to get

$$C^{1/2}\left(\int_0^\infty \cdots \int_0^\infty \|\mathcal{W}_{\lambda_{m-1}}(U_{\lambda_{m-2}}\cdots U_{\lambda_1} f)\|^2_{\mathbf{H}^1(\mathbb{R}^n)} \frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}}\right)^{1/2}$$
$$\leq 2C^{1/2}\left(\int_0^\infty \cdots \int_0^\infty \|\mathcal{W}_{\lambda_{m-1}}(U_{\lambda_{m-2}}\cdots U_{\lambda_1} f)\|^2_{\mathbf{L}^1(\mathbb{R}^n)} \frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}}\right)^{1/2}$$
$$+2C^{1/2}\sum_{j=1}^n\left(\int_0^\infty \cdots \int_0^\infty \|(R_j\mathcal{W}_{\lambda_{m-1}})(U_{\lambda_{m-2}}\cdots U_{\lambda_1} f)\|^2_{\mathbf{L}^1(\mathbb{R}^n)} \frac{d\lambda_1}{\lambda_1^{n+1}}\cdots\frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}}\right)^{1/2}$$
$$\leq (2C^{1/2}C_{m-1}^{1/2} + 2\sqrt{n}C^{1/2}C_{m-1}^{1/2})\|f\|_{\mathbb{H}^1(\mathbb{R}^n)}.$$

Thus, the theorem is proved by induction. □

3.3. $\mathbf{L}^q(\mathbb{R}^n)$ *Wavelet Scattering Transform*

We move on to the case where $1 < q < 2$.

3.3.1. *The One-layer* $\mathbf{L}^q(\mathbb{R}^n)$ *Wavelet Scattering Transform*

In this section, we prove that $S_{\text{cont},q} : \mathbf{L}^q(\mathbb{R}^n) \to \mathbf{L}^p(\mathbb{R}_+)$, where $\mathbf{L}^p(\mathbb{R}_+)$ has the weighted measure and $qp = 2$. We define:

$$\|S_{\text{cont},q} f\|^p_{\mathbf{L}^p(\mathbb{R}_+)} := \int_0^\infty |S_{\text{cont},1} f(\lambda)|^p \frac{d\lambda}{\lambda^2} = \int_0^\infty \left(\|f * \psi_\lambda\|_q^q\right)^p \frac{d\lambda}{\lambda^2},$$

and show that $\|S_{\text{cont},q} f\|_{\mathbf{L}^p(\mathbb{R}_+)} \leq C\|f\|^q_{\mathbf{L}^q(\mathbb{R}^n)}$. From this proof, it also follows that $S_{\text{dyad},q} : \mathbf{L}^q(\mathbb{R}^n) \to \ell^p(\mathbb{Z})$, where

$$\|S_{\text{dyad},q} f\|^p_{\ell^p(\mathbb{Z})} := \sum_{j\in\mathbb{Z}} |S_{\text{dyad},q} f(j)|^p = \sum_{j\in\mathbb{Z}} \left(\|f * \psi_j\|_q^q\right)^p.$$

Proposition 8. *Let $1 < q < 2$ and $qp = 2$. Also, let ψ be a wavelet that satisfies properties (24) and (25) and let $S_{cont,q}$ and $S_{dyad,q}$ be defined as above. Then there exists a universal constant $C > 0$ such that $\|S_{cont,q} f\|_{\mathbf{L}^p(\mathbb{R}_+)} \leq C\|f\|^q_{\mathbf{L}^q(\mathbb{R}^n)}$ for all $f \in \mathbf{L}^q(\mathbb{R}^n)$, and furthermore $\|S_{dyad,q} f\|_{\ell^p(\mathbb{Z})} \leq C\|f\|^q_{\mathbf{L}^q(\mathbb{R}^n)}$.*

Proof. Let $f \in \mathbf{L}^q(\mathbb{R}^n)$ with $1 < q < 2$ throughout the proof. For the continuous wavelet transform we have:

$$\|S_{\text{cont},q} f\|_{\mathbf{L}^p(\mathbb{R}_+)} = \left[\int_0^\infty \left(\|f * \psi_\lambda\|_q^q\right)^p \frac{d\lambda}{\lambda^{n+1}}\right]^{1/p}$$
$$= \left[\int_0^\infty \left(\int_{\mathbb{R}^n} |f * \psi_\lambda(x)|^q\, dx\right)^p \frac{d\lambda}{\lambda^{n+1}}\right]^{1/p}$$
$$\leq \int_{\mathbb{R}^n} \left(\int_0^\infty |f * \psi_\lambda(x)|^{qp} \frac{d\lambda}{\lambda^{n+1}}\right)^{1/p} dx$$
$$= \int_{\mathbb{R}^n} \left(\int_0^\infty |f * \psi_\lambda(x)|^2 \frac{d\lambda}{\lambda^{n+1}}\right)^{q/2} dx$$
$$= \|G(f)\|_q^q$$
$$\leq C\|f\|_q^q.$$

where in the last inequality we used Theorem 1. A similar proof follows for the dyadic case. □

3.3.2. The m-layer $\mathbf{L}^q(\mathbb{R}^n)$ Wavelet Scattering Transform

Now we try to prove that for $m \in \mathbb{N}$, $S^m_{\text{cont},q} : \mathbf{L}^q(\mathbb{R}^n) \to \mathbf{L}^p(\mathbb{R}^m_+)$. The norm for $S^m_{\text{cont},q} f$ is:

$$\begin{aligned} \|S^m_{\text{cont},q} f\|^p_{\mathbf{L}^p(\mathbb{R}^m_+)} &:= \int_0^\infty \int_0^\infty \cdots \int_0^\infty |S^m_{\text{cont},q} f(\lambda_1, \lambda_2, \ldots, \lambda_m)|^p \frac{d\lambda_1}{\lambda_1^{n+1}} \frac{d\lambda_2}{\lambda_2^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}} \\ &= \int_0^\infty \int_0^\infty \cdots \int_0^\infty \left(\|(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f) * \psi_{\lambda_m}\|^q_q \right)^p \frac{d\lambda_1}{\lambda_1^{n+1}} \frac{d\lambda_2}{\lambda_2^{n+1}} \cdots \frac{d\lambda_m}{\lambda_m^{n+1}} . \end{aligned}$$

There is also an analagous result for

$$\|S^m_{\text{dyad},q} f\|^p_{\ell^p(\mathbb{Z}^m)} := \sum_{j_m \in \mathbb{Z}} \cdots \sum_{j_m \in \mathbb{Z}} |S^m_{\text{dyad},q} f(\lambda_1, \lambda_2, \ldots, \lambda_m)|^p$$

Theorem 9. *Let $1 < q < 2$ and $qp = 2$. Also, let ψ be a wavelet that satisfies properties (24) and (25) and let $S^m_{cont,q}$ and $S^m_{dyad,q}$ be defined as above. Then there exists a universal constant $C > 0$ such that $\|S^m_{cont,q} f\|_{\mathbf{L}^p(\mathbb{R}_+)} \leq C\|f\|^q_{\mathbf{L}^q(\mathbb{R}^n)}$ for all $f \in \mathbf{L}^q(\mathbb{R}^n)$, and furthermore $\|S^m_{dyad,q} f\|_{\ell^p(\mathbb{Z})} \leq C\|f\|^q_{\mathbf{L}^q(\mathbb{R}^n)}$.*

Proof. We proceed by induction. Propositions 8 already proved the base case is satisfied. Now, let us assume that

$$\|S^{m-1}_{\text{cont},q} f\|_{\mathbf{L}^p(\mathbb{R}^{m-1}_+)} \leq C^{(m-1)/p} \|f\|^q_{\mathbf{L}^q(\mathbb{R}^n)} .$$

We now use an argument identical to Proposition 8 to get

$$\begin{aligned} \|S^{m-1}_{\text{cont},q} f\|_{\mathbf{L}^p(\mathbb{R}^m_+)} &= \left[\int_0^\infty \cdots \int_0^\infty \left(\|(U_{\lambda_{m-1}} \cdots U_{\lambda_1} f) * \psi_{\lambda_m}\|^q_q \right)^p \frac{d\lambda_m}{\lambda_m^{n+1}} \cdots \frac{d\lambda_1}{\lambda_1^{n+1}} \right]^{1/p} \\ &\leq C^{1/p} \left[\int_0^\infty \cdots \int_0^\infty \|U_{\lambda_{m-1}} \cdots U_{\lambda_1} f\|^q_q \frac{d\lambda_{m-1}}{\lambda_{m-1}^{n+1}} \cdots \frac{d\lambda_1}{\lambda_1^{n+1}} \right]^{1/p} \\ &= C^{1/p} \|S^{m-1}_{\text{cont},q} f\|_{\mathbf{L}^p(\mathbb{R}^{m-1}_+)} \\ &\leq C^{1/p} C^{(m-1)/p} \|f\|^q_q \\ &\leq C^{m/p} \|f\|^q_q . \end{aligned}$$

□

4. Stability to Dilations

We now consider dilations defined by $\tau(x) = cx$ for some constant c, so that $L_\tau f(x) = f((1-c)x)$. We will start by proving a lemma that will be useful for our work.

Lemma 10. *Assume L_τ is defined as above. Then*

$$L_\tau f * \psi_\lambda(x) = (1-c)^{-n/2} \left(f * \psi_{(1-c)\lambda} \right) ((1-c)x).$$

Proof. Notice that

$$L_\tau f * \psi_\lambda(x) = \int_{\mathbb{R}^n} f((1-c)y) \psi_\lambda(x-y)\, dy.$$

We make the substitution $z = (1-c)y$. Then it follows that

$$\begin{aligned}
L_\tau f * \psi_\lambda(x) &= (1-c)^{-n} \int_{\mathbb{R}^n} f(z)\psi_\lambda(x-(1-c)^{-1}z)\,dz \\
&= (1-c)^{-n} \int_{\mathbb{R}^n} f(z)\lambda^{-n/2}\psi\left(\lambda^{-1}(x-(1-c)^{-1}z)\right)\,dz \\
&= (1-c)^{-n/2} \int_{\mathbb{R}^n} f(z)[(1-c)\lambda]^{-n/2}\psi\left([(1-c)\lambda]^{-1}\left((1-c)x-z\right)\right)\,dz \\
&= (1-c)^{-n/2} \int_{\mathbb{R}^n} f(z)\psi_{(1-c)\lambda}\left((1-c)x-z\right)\,dz \\
&= (1-c)^{-n/2} f * \psi_{(1-c)\lambda}\left((1-c)x\right) \\
&= (1-c)^{-n/2} L_\tau\left(f * \psi_{(1-c)\lambda}\right)(x).
\end{aligned}$$

□

Remark 5. We also have

$$L_\tau \mathcal{W}_\lambda f(x) = (f * \psi_\lambda)(x(1-c)).$$

Lemma 11. *Suppose that ψ is a wavelet that satisfies the following three conditions:*

$$|\psi(x)| \le \frac{A}{(1+|x|)^{n+1+\alpha}} \quad x \in \mathbb{R}^n, \tag{34}$$

$$|\nabla\psi(x)| \le \frac{A}{(1+|x|)^{n+1+\beta}} \quad x \in \mathbb{R}^n, \tag{35}$$

$$\|D^2\psi(x)\|_\infty \le \frac{A}{(1+|x|)^{n+1+\kappa}} \quad x \in \mathbb{R}^n, \tag{36}$$

for $\alpha, \beta, \kappa > 0$. Define

$$\Psi(x) = (1-c)^{-n/2}\psi_{(1-c)}(x) - \psi(x). \tag{37}$$

Then Ψ is a wavelet satisfying (24) and (25).

Proof. Without loss of generality, assume $\alpha < \beta < \kappa < 1$. First, it's clear that $\int_{\mathbb{R}^n} \Psi = 0$. We now just need to verify properties (24) and (25). Assume $c > 0$. We can modify the proof accordingly if $c < 0$. Then

$$\begin{aligned}
|\Psi(x)| &= \left|(1-c)^{-n/2}\psi_{(1-c)}(x) - \psi(x)\right| \\
&= (1-c)^{-n}\left|\psi\left(\frac{x}{(1-c)}\right) - (1-c)^n\psi(x)\right| \\
&\le (1-c)^{-n}\left|\psi\left(\frac{x}{1-c}\right) - \psi\left(\frac{1-c}{1-c}x\right)\right| + (1-c)^{-n}\sum_{j=1}^{n}\binom{n}{j}c^j\left|\psi(x)\right|.
\end{aligned}$$

Now use mean value theorem on the first term to choose a point z on the segment connecting $\frac{x}{1-c}$ and x such that

$$\frac{c}{1-c}\left|[\nabla\psi(z)]^T x\right| = \left|\psi\left(\frac{x}{1-c}\right) - \psi\left(\frac{1-c}{1-c}x\right)\right|.$$

We now use Cauchy-Schwarz to bound the left side:

$$\frac{c}{1-c}\left|[\nabla\psi(z)]^T x\right| \le \lambda^{-1}\frac{c}{1-c}\frac{A|x|}{(1+|z|)^{n+1+\beta}}.$$

Since z lies on the segment connecting $\frac{x}{1-c}$ and x, we see that for some $t \in [0,1]$, we have

$$\begin{aligned} z &= (1-t)\frac{x}{1-c} + tx \\ &= \frac{1-t}{1-c}x + \frac{t-tc}{1-c}x \\ &= \frac{1-t+t-tc}{1-c}x \\ &= \frac{1-tc}{1-c}x. \end{aligned}$$

Thus, $|z| \geq |x|$. It now follows that

$$\frac{c}{1-c}\frac{A|x|}{(1+|z|)^{n+1+\beta}} \leq \frac{c}{1-c}\frac{A}{(1+|x|)^{n+\beta}}.$$

Finally, we get

$$\begin{aligned} |\Psi_\lambda(x)| &\leq \frac{c}{(1-c)^{n+1}}\frac{A}{(1+|x|)^{n+\beta}} + \frac{\sum_{j=1}^n \binom{n}{j} c^j}{(1-c)^{n+1}}\frac{A}{(1+|x|)^{n+\alpha}} \\ &\leq 2A\left(\frac{2n}{2n-1}\right)^{-n-1}\frac{\sum_{j=1}^n \binom{n}{j} c^j}{(1+|x|)^{n+\alpha}} \\ &\leq \frac{A_n c}{(1+|x|)^{n+\alpha}} \end{aligned}$$

for some constant A_n since we assume $\alpha < \beta$ and $c < \frac{1}{2n}$. Thus, (24) is satisfied.

We use a similar idea for proving (25) holds. Assume $c > 0$ without loss of generality. Assume $|x| \geq 2|y|$. By Mean Value Theorem, there exists z on the line segment connecting x and $x-y$ such that

$$|\Psi(x-y) - \Psi(x)| = |\nabla\Psi(z)||y|.$$

Like before, we notice that

$$\begin{aligned} |\nabla\Psi(z)| &= \left|(1-c)^{-n/2}\nabla\psi_{(1-c)}(z) - \nabla\psi(z)\right| \\ &= \left|(1-c)^{-n-1}\nabla\psi\left(\frac{z}{1-c}\right) - \nabla\psi(z)\right| \\ &= (1-c)^{-n-1}\left|\nabla\psi\left(\frac{z}{1-c}\right) - (1-c)^{n+1}\nabla\psi(z)\right| \\ &\leq (1-c)^{-n-1}\left|\nabla\psi\left(\frac{z}{1-c}\right) - \nabla\psi\left(\frac{1-c}{1-c}z\right)\right| + (1-c)^{-n-1}\sum_{j=1}^{n+1}\binom{n+1}{j}c^j\left|\nabla\psi(z)\right|. \end{aligned}$$

Let S be the set of points on the segment connecting $\frac{z}{1-c}$ and z. By Mean Value Inequality, since S is closed and bounded, we have

$$\left|\nabla\psi\left(\frac{z}{1-c}\right) - \nabla\psi\left(\frac{1-c}{1-c}z\right)\right| \leq \frac{c}{1-c}\max_{w\in S}\left\|D^2\psi(w)\right\|_\infty |z|.$$

The maximum for the quantity above is attained in S, so let's say the maximizer is $w_1 = (1-t)\frac{z}{1-c} + tz$ for some $t \in [0,1]$. Now use decay of the Hessian to bound the right side:

$$\frac{c}{1-c}\max_{w\in S}\left\|D^2\psi(w)\right\|_\infty |z| \leq \frac{c}{1-c}\frac{A|z|}{(1+|w_1|)^{n+1+\kappa}}.$$

It follows that

$$\begin{aligned}
w_1 &= (1-t)\frac{z}{1-c} + tz \\
&= \frac{1-t}{1-c}z + \frac{t-tc}{1-c}z \\
&= \frac{1-t+t-tc}{1-c}z \\
&= \frac{1-tc}{1-c}z.
\end{aligned}$$

Thus, $|w_1| \geq |z|$. We conclude that

$$\frac{c}{1-c}\frac{A|z|}{(1+|w_1|)^{n+1+\kappa}} \leq \frac{c}{1-c}\frac{A}{(1+|z|)^{n+\kappa}}.$$

For bounding $|\nabla\Psi(z)|$, we see that

$$\begin{aligned}
|\nabla\Psi(z)| &\leq \frac{c}{(1-c)^{n+2}}\frac{A}{(1+|z|)^{n+\kappa}} + \frac{\sum_{j=1}^{n+1}\binom{n+1}{j}c^j}{(1-c)^{n+1}}\frac{A}{(1+|z|)^{n+1+\beta}} \\
&\leq A(1-c)^{-n-2}\frac{2\sum_{j=1}^{n+1}\binom{n+1}{j}c^j}{(1+|z|)^{n+\kappa}} \\
&\leq \left(\frac{2n}{2n-1}\right)^{n+2}\frac{2A\sum_{j=1}^{n+1}\binom{n+1}{j}c^j}{(1+|z|)^{n+\kappa}}.
\end{aligned}$$

Going back to proving (25) holds for Ψ, we see that

$$|\Psi(x-y)-\Psi(x)| = |\nabla\Psi(z)||y| \leq \left(\frac{2n}{2n-1}\right)^{n+2}\frac{2A\sum_{j=1}^{n+1}\binom{n+1}{j}c^j|y|}{(1+|z|)^{n+\kappa}}.$$

since the point z lies on the lines on a line segment connecting $x-y$ and x with $|x| \geq 2|y|$, we can use an argument similar to above to conclude that

$$|\Psi(x-y)-\Psi(x)| \leq 2^{n+1+\kappa}\left(\frac{2n}{2n-1}\right)^{n+2}\frac{A\sum_{j=1}^{n+1}\binom{n+1}{j}c^j}{(1+|x|)^{n+\kappa}}|y|.$$

Now integrate to get

$$\begin{aligned}
\int_{|x|\geq 2|y|}|\Psi(x-y)-\Psi(x)|\,dx &\leq 2^{n+1+\kappa}\left(\frac{2n}{2n-1}\right)^{n+2} A\sum_{j=1}^{n+1}\binom{n+1}{j}c^j|y|\int_{|x|\geq 2|y|}\frac{dx}{|x|^{n+\kappa}} \\
&= 2^{n+1+\kappa}\left(\frac{2n}{2n-1}\right)^{n+2} AI_n\sum_{j=1}^{n+1}\binom{n+1}{j}c^j|y|^{1-\kappa},
\end{aligned}$$

where I_n is some constant associated with the integration. Finally, we have a bound of

$$\int_{|x|\geq 2|y|}|\Psi(x-y)-\Psi(x)|\,dx \leq \tilde{A}_n c|y|^{1-\kappa}.$$

for some constant $\tilde{A}_n$ only dependent on the dimension n. Thus, (25) holds with exponent $1-\kappa \in (0,1)$. Let $\hat{A}_n = \max\{A_n, \tilde{A}_n\}$. It follows that

$$\begin{aligned}
|\Psi_\lambda(x)| &\leq \frac{\hat{A}_n c}{(1+|x|)^{n+\alpha}} \\
\int_{|x|\geq 2|y|}|\Psi(x-y)-\Psi(x)|\,dx &\leq \hat{A}_n c|y|^{1-\kappa}.
\end{aligned}$$

$\square$

It follows from Problem 6.1.2 in [18] that the bound in the G-function depends linearly on the constant A from Theorem 1 when proving $\mathbf{L}^2(\mathbb{R}^n)$ boundedness. Thus, the following corollaries hold.

Corollary 12. *For ψ satisfying the conditions of Lemma 11, when $1 < p < \infty$, there exists a constant $C_{n,p}$ such that*

$$\left\| \left(\int_0^\infty |f * \Psi_\lambda(x)|^2 \frac{d\lambda}{\lambda^{n+1}} \right)^{1/2} \right\|_{\mathbf{L}^p(\mathbb{R}^n)} \leq c \cdot C_{n,p} \max\{p, (p-1)^{-1}\} \|f\|_{\mathbf{L}^p(\mathbb{R}^n)}$$

when $c < \frac{1}{2n}$.

Corollary 13. *For ψ satisfying the conditions of Lemma 11, there exists a constant H_n*

$$\left\| \left(\int_0^\infty |f * \Psi_\lambda(x)|^2 \frac{d\lambda}{\lambda^{n+1}} \right)^{1/2} \right\|_{\mathbf{L}^1(\mathbb{R}^n)} \leq c \cdot H_n \|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$$

when $c < \frac{1}{2n}$.

Analogous results also hold when the integral is replaced with a sum over dyadic scales.

Corollary 14. *For ψ satisfying the conditions of Lemma 11, when $1 < p < \infty$, there exists a constant $\hat{C}_{n,p}$ such that*

$$\left\| \left(\sum_{j \in \mathbb{Z}} |f * \Psi_j(x)|^2 \right)^{1/2} \right\|_{\mathbf{L}^p(\mathbb{R}^n)} \leq c \cdot \hat{C}_n \max\{p, (p-1)^{-1}\} \|f\|_{\mathbf{L}^p(\mathbb{R}^n)}$$

when $c < \frac{1}{2n}$.

Corollary 15. *For ψ satisfying the conditions of Lemma 11, there exists a constant $\hat{H}_n$ such that*

$$\left\| \left(\sum_{j \in \mathbb{Z}} |f * \Psi_j(x)|^2 \right)^{1/2} \right\|_{\mathbf{L}^1(\mathbb{R}^n)} \leq c \cdot \hat{H}_n \|f\|_{\mathbf{H}^1(\mathbb{R}^n)},$$

when $c < \frac{1}{2n}$.

4.1. One-Layer Stability to Dilations

Proposition 16. *Suppose that we consider ψ that satisfies the conditions of Lemma 11. Then there exists a constant K_n such that*

$$\|S_{cont,2} f - S_{cont,2} L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \leq c \cdot K_n \|f\|_2^2$$

for any $c < \frac{1}{2n}$.

Proof. We use the following notation throughout this proposition:

$$\begin{aligned} \mathcal{W} f &= (f * \psi_\lambda)_{\lambda \in (0,\infty)} \\ Mf &= |f| \\ A_2 f &= \int_{\mathbb{R}^n} f^2(x)\, dx. \end{aligned}$$

Then we have

$$
\begin{aligned}
\|S_{\text{cont},2}f - S_{\text{cont},2}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} &= \|A_2 M\mathcal{W}f - A_2 M\mathcal{W}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \\
&= \|A_2 M\mathcal{W}f - A_2 ML_\tau\mathcal{W}f + A_q ML_\tau\mathcal{W}f - A_q M\mathcal{W}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \\
&\leq \|A_2 M\mathcal{W}f - A_2 ML_\tau\mathcal{W}f\|_{\mathbf{L}^1(\mathbb{R}_+)} + \|A_2 ML_\tau\mathcal{W}f - A_2 M\mathcal{W}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \\
&= \|A_2 M\mathcal{W}f - A_2 ML_\tau\mathcal{W}f\|_{\mathbf{L}^1(\mathbb{R}_+)} + \|A_2 ML_\tau\mathcal{W}f - A_2 M\mathcal{W}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \\
&= \|A_2 M\mathcal{W}f - A_2 ML_\tau\mathcal{W}f\|_{\mathbf{L}^1(\mathbb{R}_+)} + \|A_2 M[\mathcal{W}, L_\tau]f\|_{\mathbf{L}^1(\mathbb{R}_+)}.
\end{aligned}
$$

We'll bound the first term in this sum first. We have

$$
|A_2 M\mathcal{W}f - A_2 ML_\tau\mathcal{W}f| = \left| \|\mathcal{W}f\|_2^2 - \|\mathcal{W}L_\tau f\|_2^2 \right|.
$$

Make the substitution $z = (1-c)x$. Then

$$
\begin{aligned}
\|\mathcal{W}L_\tau f\|_2^2 &= \int_{\mathbb{R}^n} |\mathcal{W}f((1-c)x)|^2\, dx \\
&= \frac{1}{(1-c)^n} \int_{\mathbb{R}^n} |\mathcal{W}f(z)|^2\, dz \\
&= \frac{1}{(1-c)^n} \|\mathcal{W}f\|_2^2.
\end{aligned}
$$

Thus, it follows that

$$
\begin{aligned}
|A_2 M\mathcal{W}f - A_2 ML_\tau\mathcal{W}f| &= \|\mathcal{W}f\|_2^2 \left(1 - \frac{1}{(1-c)^n}\right) \\
&= \frac{\sum_{j=1}^n \binom{n}{j} c^j}{(1-c)^n} \|\mathcal{W}f\|_2^2 \\
&\leq \left(\frac{2n}{2n-1}\right)^n \sum_{j=1}^n \binom{n}{j} c^j \|\mathcal{W}f\|_2^2.
\end{aligned}
$$

Using the result above,

$$
\begin{aligned}
\|A_2 M\mathcal{W}f - A_2 ML_\tau\mathcal{W}f\|_{L^1(\mathbb{R}_+)} &= \int_0^\infty |(A_2 M - A_2 ML_\tau)\mathcal{W}f(\lambda)| \frac{d\lambda}{\lambda^{n+1}} \\
&\leq \int_0^\infty \left(\frac{2n}{2n-1}\right)^n \sum_{j=1}^n \binom{n}{j} c^j \|f * \psi_\lambda\|_2^2 \frac{d\lambda}{\lambda^{n+1}} \\
&\leq c \cdot B_n \|f\|_2^2.
\end{aligned}
$$

for some constant B_n. This bounds the first term appropriately.

Now, we move on to the second term. We have

$$
\begin{aligned}
\|A_2 M[\mathcal{W}, L_\tau]f\|_{\mathbf{L}^1(\mathbb{R}_+)} &= \left\| \|[\mathcal{W}, L_\tau]f\|_2^2 \right\|_{\mathbf{L}^1(\mathbb{R}_+)} \\
&= \int_0^\infty \|[\mathcal{W}, L_\tau]f\|_2^2 \frac{d\lambda}{\lambda^{n+1}} \\
&= \|[\mathcal{W}, L_\tau]f\|_{\mathbf{L}^1(\mathbb{R}_+, \mathbf{L}^2(\mathbb{R}^n))}^2.
\end{aligned}
$$

Thus, once we bound this quantity appropriately, we'll finish the proof. We start by writing

$$
\|[\mathcal{W}, L_\tau]f\|_{\mathbf{L}^1(\mathbb{R}_+, \mathbf{L}^2(\mathbb{R}^n))} = \int_0^\infty \|(L_\tau f) * \psi_\lambda - L_\tau (f * \psi_\lambda)\|_2^2 \frac{d\lambda}{\lambda^{n+1}}.
$$

By substitution with $z = (1-c)x$ and Lemma 10,

$$\begin{aligned}
\|(L_\tau f) * \psi_\lambda - L_\tau (f * \psi_\lambda)\|_2^2 &= \int_{\mathbb{R}^n} |(L_\tau f * \psi_\lambda)(x) - L_\tau (f * \psi_\lambda)(x)|^2 \, dx \\
&= \int_{\mathbb{R}^n} \left|(1-c)^{-n/2} \left(f * \psi_{(1-c)\lambda}\right)((1-c)x) - (f * \psi_\lambda)((1-c)x)\right|^2 dx \\
&= (1-c)^{-n} \int_{\mathbb{R}^n} \left|(1-c)^{-n/2} \left(f * \psi_{(1-c)\lambda}\right)(z) - (f * \psi_\lambda)(z)\right|^2 dz \\
&= (1-c)^{-n} \int_{\mathbb{R}^n} \left|f * \left((1-c)^{-n/2} \psi_{(1-c)\lambda} - \psi_\lambda\right)\right|^2 dz \\
&= (1-c)^{-n} \int_{\mathbb{R}^n} |(f * \Psi_\lambda)(z)|^2 \, dz, \\
&= (1-c)^{-n} \|f * \Psi_\lambda\|_2^2.
\end{aligned}$$

We then obtain

$$\begin{aligned}
\int_0^\infty \|(L_\tau f) * \psi_\lambda - L_\tau (f * \psi_\lambda)\|_2^2 \frac{d\lambda}{\lambda^{n+1}} &= (1-c)^{-n} \int_0^\infty \|f * \Psi_\lambda\|_2^2 \frac{d\lambda}{\lambda^{n+1}} \\
&= (1-c)^{-n} \int_{\mathbb{R}^n} \int_0^\infty |f * \Psi_\lambda(x)|^2 \frac{d\lambda}{\lambda^{n+1}} \, dx \\
&= (1-c)^{-n} \left\| \left(\int_0^\infty |f * \Psi_\lambda(x)|^2 \frac{d\lambda}{\lambda^{n+1}} \right) \right\|_2^2 \\
&\leq c \cdot \left(\frac{2n}{2n-1}\right)^n C_{n,p} \|f\|_2^2 \\
&\leq c \cdot \tilde{C}_n \|f\|_2^2.
\end{aligned}$$

It now follows that

$$\|S_{\text{cont},2} f - S_{\text{cont},2} L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \leq c(B_n + \tilde{C}_n)\|f\|_2^2.$$

Let $K_n := B_n + \tilde{C}_n$. This finishes the proof of stability dilation. □

In the case where $1 \leq q < 2$, we can use a similar argument to bound each term. The main difference is in bounding the commutator. However, bounding the commutator when $1 \leq q < 2$ follows from extending by means of Theorem 1 since we proved a bound when $q = 2$. Thus, the following results hold for the continuous transforms.

Corollary 17. *For $q \in (1,2)$, there exists a constant $K_{n,q}$ such that*

$$\|S_{cont,q} f - S_{cont,q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}_+)} \leq c \cdot K_{n,q} \|f\|_q^q$$

for any $c < \frac{1}{2n}$.

Corollary 18. *Additionally assume that ψ has $n+3$ vanishing moments in one direction. There exists a constant K_H such that*

$$\|S_{cont,1} f - S_{cont,1} L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \leq c \cdot H_{n,m} \|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$$

for any $c < \frac{1}{2n}$.

We also have the following bounds that hold for the dyadic transform. The proof is almost identical other than switching the integral to a sum.

Corollary 19. *There exists a constant $\hat{K}_n$ such that*

$$\|S_{dyad,2}f - S_{dyad,2}L_\tau f\|_{\ell^2(\mathbb{Z})} \leq c \cdot \hat{K}_n \|f\|_2^2$$

for any $c < \frac{1}{2n}$.

Corollary 20. *There exists a constant $\hat{K}_{n,p}$ such that*

$$\|S_{dyad,q}f - S_{dyad,q}L_\tau f\|_{\ell^p(\mathbb{Z})} \leq c \cdot \hat{K}_{n,p} \|f\|_q^q$$

for any $c < \frac{1}{2n}$.

Corollary 21. *There exists a constant $\hat{K}_H$ such that*

$$\|S_{dyad,1}f - S_{dyad,1}L_\tau f\|_{\ell^2(\mathbb{Z})} \leq c \cdot \hat{H}_{n,m} \|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$$

for any $c < \frac{1}{2n}$.

4.2. m-Layer Stability to Dilations

Theorem 22. *Suppose that ψ is a wavelet that satisfies the conditions of Lemma 11. Then there exists a constant $K_{n,m}$ only dependent on n and m such that*

$$\|S^m_{cont,2}f - S^m_{cont,2}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}_+)} \leq c \cdot K_{n,m} \|f\|_2^2$$

for any $c < \frac{1}{2n}$.

Proof. Use the same notation as in previous propositions. Define the operator $U_t = M\mathcal{W}_t$. It follows that $S^m_{\text{dyad},2} = A_2 M W_{\lambda_m} U_{\lambda_{m-1}} \cdots U_{\lambda_1}$. We will let $V_{m-1} = U_{\lambda_{m-1}} \cdots U_{\lambda_1}$ and make a slight abuse of notation by denoting $\mathcal{W}_{\lambda_m}$ as $\mathcal{W}$ since the proof will reduce to bounding the commutator like in the one layer case. First, we have

$$\begin{aligned}
\|S^m_{\text{cont},2}f - S^m_{\text{cont},2}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} &= \|A_2 M\mathcal{W}V_{m-1}f - A_2 M\mathcal{W}V_{m-1}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \\
&= \|A_2 M\mathcal{W}V_{m-1}f - A_2 ML_\tau\mathcal{W}V_{m-1}f + A_2 ML_\tau\mathcal{W}V_{m-1}f - A_2 M\mathcal{W}V_{m-1}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \\
&\leq \|A_2 M\mathcal{W}V_{m-1}f - A_2 ML_\tau\mathcal{W}V_{m-1}f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} + \|A_2 M[\mathcal{W}V_{m-1}, L_\tau]f\|_{\mathbf{L}^1(\mathbb{R}^m_+)}.
\end{aligned}$$

Like before, we'll start by bounding the first term. We see that $g = \mathcal{W}V_{m-1}f \in \mathbf{L}^2(\mathbb{R}^n)$. Thus

$$|A_2 M\mathcal{W}V_{m-1}f - A_2 ML_\tau\mathcal{W}V_{m-1}f| = \left| \|g\|_2^2 - \|L_\tau g\|_2^2 \right|.$$

Now use a change of variables like before:

$$\|L_\tau g\|_2^2 = \int_{\mathbb{R}^n} |g((1-c)x)|^2 \, dx = (1-c)^{-n} \|g\|_2^2.$$

It then follows that

$$\left| \|g\|_2^2 - \|L_\tau g\|_2^2 \right| = \|g\|_2^2 \left(\frac{1}{(1-c)^n} - 1 \right) = \frac{1}{(1-c)^n} \sum_{j=1}^n \binom{n}{j} c^j \|g\|_2^2 \leq \left(\frac{2n}{2n-1} \right)^n \sum_{j=1}^n \binom{n}{j} c^j \|g\|_2^2.$$

Now take the norm to get

$$\begin{aligned}
\|A_2 M\mathcal{W}V_{m-1}f - A_2 ML_\tau\mathcal{W}V_{m-1}f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} &\leq \left(\frac{2n}{2n-1} \right)^n \sum_{j=1}^n \binom{n}{j} c^j \|S^m_{\text{dyad},2}f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \\
&\leq c \cdot C_{m,n} \|f\|_2.
\end{aligned}$$

For the second term,

$$\|A_2 M[\mathcal{W}U_{m-1}, L_\tau]f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \leq \|[\mathcal{W}U_{m-1}, L_\tau]\|^2 \|f\|_2^2.$$

We examine the commutator term more closely. By expanding it, we see that each term contains $[\mathcal{W}, L_\tau]$. It follows that

$$\|[\mathcal{W}U_{m-1}, L_\tau]\| \leq m\|\mathcal{W}\|^{m-1}\|M\|^{m-1}\|[\mathcal{W}, L_\tau]\| \leq C_m\|[\mathcal{W}, L_\tau]\|.$$

We proved this term was already bounded, so it follows that

$$\|S^m_{\text{cont},2}f - S^m_{\text{cont},2}L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \leq c \cdot K_{n,m}\|f\|_2^2$$

for any $c < \frac{1}{2n}$. □

As is customary at this point, we have the following corollaries.

Corollary 23. *For $q \in (1, 2)$, there exists a constant $K_{n,m,q}$ such that*

$$\|S^m_{cont,q}f - S^m_{cont,q}L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)} \leq c \cdot K_{n,m,q}\|f\|_q^q$$

for any $c < \frac{1}{2n}$.

Corollary 24. *Additionally assume that ψ has $n+3$ vanishing moments in one direction. There exists a constant $K_{H,m}$ such that*

$$\|S^m_{cont,1}f - S^m_{cont,1}L_\tau f\|_{\mathbf{L}^2(\mathbb{R}^m_+)} \leq c \cdot K_{H,m}\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$$

for any $c < \frac{1}{2n}$.

Corollary 25. *There exists a constant $\hat{K}_{n,m}$ such that*

$$\|S^m_{dyad,2}f - S^m_{dyad,2}L_\tau f\|_{\ell^2(\mathbb{Z}^m)} \leq c \cdot \hat{K}_{n,m}\|f\|_2^2$$

for any $c < \frac{1}{2n}$.

Corollary 26. *There exists a constant $\hat{K}_{n,q}$ such that*

$$\|S^m_{dyad,q}f - S^m_{dyad,q}L_\tau f\|_{\ell^p(\mathbb{Z}^m)} \leq c \cdot \hat{K}_{n,m,q}\|f\|_q^q$$

for any $c < \frac{1}{2n}$.

Corollary 27. *There exists a constant $\hat{K}_{H,m}$ such that*

$$\|S^m_{dyad,1}f - S^m_{dyad,1}L_\tau f\|_{\ell^2(\mathbb{Z}^m)} \leq c \cdot \hat{K}_{H,m}\|f\|_{\mathbf{H}^1(\mathbb{R}^n)}$$

for any $c < \frac{1}{2n}$.

5. Stability to Diffeomorphisms

We now focus on the stability of $S^m_{\text{cont},q}f$ for general diffeomorphisms with $\|D\tau\|_\infty < \frac{1}{2n}$. The corresponding operator for diffeomorphisms is defined as $L_\tau f(x) = f(x - \tau(x))$.

5.1. Stability to Diffeomorphisms When $q = 2$

Lemma 28. *Define $\phi_t(x) = t^{n/2}\psi(tx)$. We have*

$$\|[\mathcal{W}, L_\tau]f\|_2^2 = \int_0^\infty \int_{\mathbb{R}^n} |(\phi_t * L_\tau f)(x) - L_\tau(\phi_t * f)(x)|^2 \, dx \, t^{n-1} \, dt.$$

Proof. The proof follows by a straightforward calculation:

$$\begin{aligned}
\|[\mathcal{W}, L_\tau]f\|_2^2 &= \int_0^\infty \|[\mathcal{W}_\lambda, L_\tau]f\|_2^2 \frac{d\lambda}{\lambda^{n+1}} \\
&= \int_0^\infty \|\psi_\lambda * (L_\tau f) - L_\tau(\psi_\lambda * f)\|_2^2 \frac{d\lambda}{\lambda^{n+1}}.
\end{aligned}$$

Now notice that $\psi_{\frac{1}{t}}(x) = t^{-n/2}\psi(tx) = \phi_t(x)$. Let $\lambda = \frac{1}{t}$. Then we have

$$\begin{aligned}
\|[\mathcal{W}, L_\tau]f\|_2^2 &= \int_0^\infty \|\psi_{\frac{1}{t}} * (L_\tau f) - L_\tau(\psi_{\frac{1}{t}} * f)\|_2^2 t^{n-1} \, dt \\
&= \int_0^\infty \|\phi_t * (L_\tau f) - L_\tau(\phi_t * f)\|_2^2 \, t^{n-1} \, dt.
\end{aligned}$$

□

We can now switch normalizations for the next proposition. Define $\psi_\lambda(x) = \lambda^{n/2}\psi(\lambda x)$. Then define the operator $\mathcal{W}_\lambda f = \psi_\lambda * f$. We will bound

$$\|[\mathcal{W}, L_\tau]f\|_2^2 := \int_0^\infty \|[\mathcal{W}_\lambda, L_\tau]f\|_2^2 \, \lambda^{n-1} \, d\lambda.$$

using our new normalization, and it will be equivalent to bounding the commutator with our previous normalization because of the lemma above.

Proposition 29. *With the normalization defined above, assume ψ and its first and second order derivatives have decay in $O((1 + |x|)^{-n-2})$, and $\int_{\mathbb{R}^n} \psi(x) \, dx = 0$. Then for every $\tau \in \mathbf{C}^2(\mathbb{R}^n)$ with $\|D\tau\|_\infty \leq \frac{1}{2n}$, there exists $\tilde{C}_n > 0$ such that:*

$$\|[\mathcal{W}, L_\tau]\| \leq \tilde{C}_n \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right).$$

Proof. We outline the proof and show how it generalizes to the continuous case. First, we have

$$\|[\mathcal{W}, L_\tau]f\|_2^2 = \int_0^\infty \|[\mathcal{W}_\lambda, L_\tau]f\|_2^2 \, \lambda^{n-1} \, d\lambda.$$

This implies

$$[\mathcal{W}, L_\tau]^*[\mathcal{W}, L_\tau] = \int_0^\infty [\mathcal{W}_\lambda, L_\tau]^*[\mathcal{W}_\lambda, L_\tau] \, \lambda^{n-1} \, d\lambda.$$

Define $Z_\lambda = f * \lambda^{n/2}\psi_\lambda$, i.e. convolution with an $\mathbf{L}^1$ normalized wavelet so that $\mathcal{W}_\lambda = \lambda^{-n/2} Z_\lambda$. Defining

$K_\lambda = Z_\lambda - L_\tau Z_\lambda L_\tau^{-1}$ so that $[Z_\lambda, L_\tau] = K_\lambda L_\tau$, we have:

$$\begin{aligned}
\|[\mathcal{W}, L_\tau]\| &= \|[\mathcal{W}, L_\tau]^*[\mathcal{W}, L_\tau]\|^{1/2} \\
&= \left\| \int_0^\infty [\mathcal{W}_\lambda, L_\tau]^*[\mathcal{W}_\lambda, L_\tau]\, \lambda^{n-1}\, d\lambda \right\|^{1/2} \\
&= \left\| \int_0^\infty [Z_\lambda, L_\tau]^*[Z_\lambda, L_\tau]\, \frac{d\lambda}{\lambda} \right\|^{1/2} \\
&= \left\| \int_0^\infty L_\tau^* K_\lambda^* K_\lambda L_\tau\, \frac{d\lambda}{\lambda} \right\|^{1/2} \\
&\le \|L_\tau\| \cdot \left\| \int_0^\infty K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\|^{1/2}.
\end{aligned}$$

Since $\|L_\tau f\|_2^2 \le \left(\frac{1}{1-n\|\tau\|_\infty}\right)\|f\|_2^2$,

$$\|L_\tau\| \le \frac{1}{1 - n\|D\tau\|_\infty} \le 2$$

and the problem is reduced to bounding $\left\|\int_0^\infty K_\lambda^* K_\lambda\, \lambda^{-1}\, d\lambda\right\|^{1/2}$. The integral is divided into three pieces:

$$\begin{aligned}
\left\| \int_0^\infty K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\|^{1/2} &\le \left(\left\| \int_0^{2^{-\gamma}} K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\| + \left\| \int_{2^{-\gamma}}^1 K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\| + \left\| \int_1^\infty K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\| \right)^{1/2} \\
&\le \left\| \int_0^{2^{-\gamma}} K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\|^{1/2} + \left\| \int_{2^{-\gamma}}^1 K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\|^{1/2} + \left\| \int_1^\infty K_\lambda^* K_\lambda\, \frac{d\lambda}{\lambda} \right\|^{1/2} \\
&= P_1 + P_2 + P_3.
\end{aligned}$$

To bound P_1, we decompose $K_\lambda = \tilde{K}_{\lambda,1} + \tilde{K}_{\lambda,2}$, where the kernels defining $\tilde{K}_{\lambda,1}, \tilde{K}_{\lambda,2}$ are

$$\begin{aligned}
\tilde{k}_{\lambda,1}(x,u) &:= (1 - \det(I - D\tau(u)))\lambda^{n/2}\psi_\lambda(x-u), \\
\tilde{k}_{\lambda,2}(x,u) &:= \det(I - D\tau(u))(\lambda^{n/2}\psi_\lambda(x-u) - \lambda^{n/2}\psi_\lambda(x - \tau(x) - u + \tau(u))),
\end{aligned}$$

respectively. In [5] it is shown that there exists a constant C_n such that

$$\begin{aligned}
\|\tilde{K}_{\lambda,2}\| &\le C_n \lambda \|\Delta\tau\|_\infty \\
\|\tilde{K}_{\lambda,1}\| &\le C_n \|D\tau\|_\infty.
\end{aligned}$$

We want to prove that

$$\left\| \int_0^1 \tilde{K}_{\lambda,1}^* \tilde{K}_{\lambda,1}\, \frac{d\lambda}{\lambda} \right\|^{1/2} \le C_n \|D\tau\|_\infty,$$

and the steps in [5] imply it suffices to prove that the function

$$\theta(x) = \int_0^1 \psi_\lambda * \psi_\lambda(x) d\lambda$$

satisfies $\theta \in \mathbf{L}^1(\mathbb{R}^n)$. To prove this, we simply bound:

$$\begin{aligned}
\int_{\mathbb{R}^n} |\theta(x)|\, dx &= \int_{\mathbb{R}^n} \left| \int_0^1 \psi_\lambda * \psi_\lambda(x) d\lambda \right| dx \\
&\leq \int_{\mathbb{R}^n} \int_0^1 |\psi_\lambda * \psi_\lambda(x)|^2 d\lambda\, dx \\
&= \int_0^1 \int_{\mathbb{R}^n} |\psi_\lambda * \psi_\lambda(x)|^2 dx\, d\lambda \\
&\leq \int_0^1 \|\psi_\lambda\|_1^2 d\lambda \\
&= \|\psi\|_1^2 < \infty
\end{aligned}$$

because of our wavelet normalization. Thus, it follows that

$$\begin{aligned}
\left\| \int_0^{2^{-\gamma}} K_\lambda^* K_\lambda \, \frac{d\lambda}{\lambda} \right\|^{1/2} &= \left\| \int_0^{2^{-\gamma}} (\tilde{K}_{\lambda,1} + \tilde{K}_{\lambda,2})^* (\tilde{K}_{\lambda,1} + \tilde{K}_{\lambda,2}) \, \frac{d\lambda}{\lambda} \right\|^{1/2} \\
&= \left\| \int_0^{2^{-\gamma}} (\tilde{K}_{\lambda,1}^* \tilde{K}_{\lambda,1} + \tilde{K}_{\lambda,1}^* \tilde{K}_{\lambda,2} + \tilde{K}_{\lambda,2}^* \tilde{K}_{\lambda,1} + \tilde{K}_{\lambda,2}^* \tilde{K}_{\lambda,2}) \, \frac{d\lambda}{\lambda} \right\|^{1/2} \\
&\leq \left(\left\| \int_0^{2^{-\gamma}} \tilde{K}_{\lambda,1}^* \tilde{K}_{\lambda,1} \frac{d\lambda}{\lambda} \right\| + \left\| \int_0^{2^{-\gamma}} \tilde{K}_{\lambda,2}^* \tilde{K}_{\lambda,2} + \tilde{K}_{\lambda,2}^* \tilde{K}_{\lambda,1} + \tilde{K}_{\lambda,2}^* \tilde{K}_{\lambda,2} \, \frac{d\lambda}{\lambda} \right\| \right)^{1/2} \\
&\leq \left(\left\| \int_0^{2^{-\gamma}} \tilde{K}_{\lambda,1}^* \tilde{K}_{\lambda,1} \frac{d\lambda}{\lambda} \right\| + \int_0^{2^{-\gamma}} \|\tilde{K}_{\lambda,2}\|^2 \, \frac{d\lambda}{\lambda} + \int_0^{2^{-\gamma}} 2\|\tilde{K}_{\lambda,1}\| \|\tilde{K}_{\lambda,2}\| \, \frac{d\lambda}{\lambda} \right)^{1/2} \\
&\leq \left\| \int_0^{2^{-\gamma}} \tilde{K}_{\lambda,1}^* \tilde{K}_{\lambda,1} \frac{d\lambda}{\lambda} \right\|^{1/2} + \left(\int_0^{2^{-\gamma}} \|\tilde{K}_{\lambda,2}\|^2 \, \frac{d\lambda}{\lambda} \right)^{1/2} + \left(\int_0^{2^{-\gamma}} 2\|\tilde{K}_{\lambda,1}\| \|\tilde{K}_{\lambda,2}\| \, \frac{d\lambda}{\lambda} \right)^{1/2} \\
&\leq 2C_n \left(\|D\tau\|_\infty + \|\Delta\tau\|_\infty \left(\int_0^{2^{-\gamma}} \lambda^2 \, \frac{d\lambda}{\lambda} \right)^{1/2} + \|D\tau\|_\infty^{1/2} \|\Delta\tau\|_\infty^{1/2} \left(\int_0^{2^{-\gamma}} 2\lambda \, \frac{d\lambda}{\lambda} \right)^{1/2} \right) \\
&\leq 2C_n \left(\|D\tau\|_\infty + 2^{-\gamma} \|\Delta\tau\|_\infty + 2^{-\gamma/2} \|D\tau\|_\infty^{1/2} \|\Delta\tau\|_\infty^{1/2} \right) \\
&\leq 4C_n \left(\|D\tau\|_\infty + 2^{-\gamma} \|\Delta\tau\|_\infty \right).
\end{aligned}$$

To bound P_3, we decompose $K_\lambda = K_{\lambda,1} + K_{\lambda,2}$, where the kernels defining $K_{\lambda,1}, K_{\lambda,2}$ are

$$\begin{aligned}
k_{\lambda,1}(x,u) &= \lambda^{n/2} \psi_\lambda(x-u) - \lambda^{n/2} \psi_\lambda((I - D\tau(u))(x-u)) \det(I - D\tau(u)) \\
k_{\lambda,2}(x,u) &= \det(I - D\tau(u)) \lambda^{n/2} \psi_\lambda((I - D\tau(u))(x-u)) - \lambda^{n/2} \psi_\lambda(x - \tau(x) - u + \tau(u))).
\end{aligned}$$

A similar computation to the one for P_1 shows that:

$$\left\| \int_1^\infty K_\lambda^* K_\lambda \, \frac{d\lambda}{\lambda} \right\|^{1/2} \leq \left\| \int_1^\infty K_{\lambda,1}^* K_{\lambda,1} \frac{d\lambda}{\lambda} \right\|^{1/2} + \left(\int_1^\infty \|K_{\lambda,2}\|^2 \, \frac{d\lambda}{\lambda} \right)^{1/2} + \left(\int_1^\infty 2\|K_{\lambda,1}\| \|K_{\lambda,2}\| \, \frac{d\lambda}{\lambda} \right)^{1/2}$$

Letting $Q_j = K_{2^j,1}^* K_{2^j,1}$, it is shown in [5] that:

$$\begin{aligned}
\|K_{\lambda,1}\| &\leq C_n \|D\tau\|_\infty \\
\|K_{\lambda,2}\| &\leq \min\{\lambda^{-n} \|D^2\tau\|_\infty, \|D\tau\|_\infty\} \\
\|Q_j Q_\ell\| &\leq C_n^2 2^{-|j-\ell|} (\|D\tau\|_\infty + \|D^2\tau\|_\infty)^4
\end{aligned}$$

so that

$$\left\| \int_1^\infty K_{\lambda,1}^* K_{\lambda,1} \frac{d\lambda}{\lambda} \right\|^{1/2} = \left\| \int_0^\infty K_{2^j,1}^* K_{2^j,1} \log(2)\, dj \right\|^{1/2}$$
$$= \sqrt{\log(2)} \left\| \int_0^\infty Q_j\, dj \right\|^{1/2}.$$

We now apply a continuous version of Cotlar's Lemma (see Ch. 7 of [20], Sec. 5.5 for the continuous extension). We define:

$$\beta(j,\ell) = \begin{cases} C_n 2^{-|j-\ell|/2} (\|D\tau\|_\infty + \|D^2\tau\|_\infty)^2 & j \geq 0 \text{ and } \ell \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Defining $Q_j = 0$ for $j < 0$, we have $\|Q_j^* Q_\ell\| \leq \beta(j,\ell)^2$ and $\|Q_j Q_\ell^*\| \leq \beta(j,\ell)^2$ for all j, ℓ. Thus by Cotlar's Lemma:

$$\left\| \int_{\mathbb{R}} Q_j\, dj \right\| \leq \sup_{j \in \mathbb{R}} \int_{\mathbb{R}} \beta(j,\ell)\, d\ell,$$
$$\left\| \int_0^\infty Q_j\, dj \right\| \leq \sup_{j \geq 0} \int_0^\infty \beta(j,\ell)\, d\ell$$
$$\leq C_n (\|D\tau\|_\infty + \|H\tau\|_\infty)^2 \left(\sup_{j \geq 0} \int_0^\infty 2^{-|j-\ell|/2}\, d\ell \right).$$

Now observing that with the change of variable $\lambda_1 = 2^j, \lambda_2 = 2^\ell$, we have $2^{-|j-\ell|/2} = \frac{\lambda_1}{\lambda_2} \wedge \frac{\lambda_2}{\lambda_1}$, we obtain:

$$\begin{aligned}
\sup_{j \geq 0} \int_0^\infty 2^{-|j-\ell|/2}\, d\ell &= \sup_{\lambda_1 \geq 1} \int_1^\infty \frac{(\lambda_1 \wedge \lambda_2)}{\sqrt{\lambda_1 \lambda_2}} \frac{d\lambda_2}{\ln(2)\lambda_2} \\
&= \frac{1}{\ln(2)} \sup_{\lambda_1 \geq 1} \left(\int_1^{\lambda_1} \frac{1}{\sqrt{\lambda_1 \lambda_2}}\, d\lambda_2 + \int_{\lambda_1}^\infty \frac{\sqrt{\lambda_1}}{\lambda_2^{3/2}}\, d\lambda_2 \right) \\
&= \frac{1}{\ln(2)} \sup_{\lambda_1 \geq 1} \left(\frac{1}{\sqrt{\lambda_1}} (2\sqrt{\lambda_1} - 2) + \sqrt{\lambda_1} \left(\frac{2}{\sqrt{\lambda_1}} \right) \right) \\
&= \frac{1}{\ln(2)} \sup_{\lambda_1 \geq 1} \left(4 - \frac{2}{\sqrt{\lambda_1}} \right) \\
&= \frac{4}{\ln(2)}
\end{aligned}$$

and conclude that

$$\left\| \int_1^\infty K_{\lambda,1}^* K_{\lambda,1} \frac{d\lambda}{\lambda} \right\|^{1/2} \leq 3C_n (\|D\tau\|_\infty + \|H\tau\|_\infty).$$

Thus we have:

$$\left\| \int_1^\infty K_\lambda^* K_\lambda \frac{d\lambda}{\lambda} \right\|^{1/2} \leq \left\| \int_1^\infty K_{\lambda,1}^* K_{\lambda,1} \frac{d\lambda}{\lambda} \right\|^{1/2} + \left(\int_1^\infty \|K_{\lambda,2}\|^2 \frac{d\lambda}{\lambda} \right)^{1/2} + \left(\int_1^\infty 2\|K_{\lambda,1}\| \|K_{\lambda,2}\| \frac{d\lambda}{\lambda} \right)^{1/2}.$$

Now we see that there exists a constant C_n such that

$$\left\| \int_1^\infty K_{\lambda,1}^* K_{\lambda,1} \frac{d\lambda}{\lambda} \right\|^{1/2} \le C_n(\|D\tau\|_\infty + \|D^2\tau\|_\infty)$$
$$\left(\int_1^\infty \|K_{\lambda,2}\|^2 \frac{d\lambda}{\lambda} \right)^{1/2} \le C_n \|D^2\tau\|_\infty \left(\int_1^\infty \lambda^{-2n} \frac{d\lambda}{\lambda} \right)^{1/2}$$
$$\left(\int_1^\infty 2\|K_{\lambda,1}\|\|K_{\lambda,2}\| \frac{d\lambda}{\lambda} \right)^{1/2} \le C_n \|D\tau\|_\infty^{1/2} \|D^2\tau\|_\infty^{1/2} \left(\int_1^\infty 2\lambda^{-n} \frac{d\lambda}{\lambda} \right)^{1/2}.$$

Thus, we have

$$\begin{aligned} \left\| \int_1^\infty K_\lambda^* K_\lambda \frac{d\lambda}{\lambda} \right\|^{1/2} &\le C_n \left(\|D\tau\|_\infty + \frac{1}{2n}\|D^2\tau\|_\infty + \frac{2}{n}\|D\tau\|_\infty^{1/2}\|D^2\tau\|_\infty^{1/2} \right) \\ &\le C_n \left(\|D\tau\|_\infty + \frac{1}{2n}\|D^2\tau\|_\infty + \frac{1}{n}\|D\tau\|_\infty + \frac{1}{n}\|D^2\tau\|_\infty \right) \\ &\le 2C_n(\|D\tau\|_\infty + \|D^2\tau\|_\infty). \end{aligned}$$

Finally, we bound P_2. Note that in the previous section it was observed (shown in [5]) that

$$\begin{aligned} &\|K_{\lambda,1}\| \le C_n \|D\tau\|_\infty \\ &\|K_{\lambda,2}\| \le \min\{\lambda^{-n}\|D^2\tau\|_\infty, \|D\tau\|_\infty\}. \end{aligned}$$

The above two inequalities imply

$$\|K_\lambda\| = \|K_{\lambda,1} + K_{\lambda,2}\| \le \|K_{\lambda,1}\| + \|K_{\lambda,2}\| \le 2C_n\|D\tau\|_\infty$$

so that

$$\begin{aligned} \left\| \int_{2^{-\gamma}}^1 K_\lambda^* K_\lambda \frac{d\lambda}{\lambda} \right\|^{1/2} &\le \left(\int_{2^{-\gamma}}^1 \|K_\lambda\|^2 \frac{d\lambda}{\lambda} \right)^{1/2} \\ &\le 2C_n\|D\tau\|_\infty \left(\int_{2^{-\gamma}}^1 \frac{d\lambda}{\lambda} \right)^{1/2} \\ &\le 2C_n\|D\tau\|_\infty \left(-\ln(2^{-\gamma})\right)^{1/2} \\ &\le 2C_n\gamma^{1/2}\|D\tau\|_\infty. \end{aligned}$$

Putting everything together and assuming $\gamma \ge 1$, we obtain:

$$\begin{aligned} \|[\mathcal{W}, L_\tau]\| &\le 2(P_1 + P_2 + P_3) \\ &\le 4C_n \left(\|D\tau\|_\infty + 2^{-\gamma}\|\Delta\tau\|_\infty\right) + 2C_n\gamma^{1/2}\|D\tau\|_\infty + 3C_n(\|D\tau\|_\infty + \|D^2\tau\|_\infty) \\ &\le \tilde{C}_n \left(\gamma\|D\tau\|_\infty + 2^{-\gamma}\|\Delta\tau\|_\infty + \|D^2\tau\|_\infty \right). \end{aligned}$$

Choosing $\gamma = \left(\log \frac{\|\Delta\tau\|_\infty}{\|\tau'\|_\infty}\right) \vee 1$ gives

$$\|[\mathcal{W}, L_\tau]\| \le \tilde{C}_n \left(\left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) \|D\tau\|_\infty + \|D^2\tau\|_\infty \right),$$

and the lemma is proved. □

Corollary 30. *Under the conditions given in the previous theorem, if we use our previous normalization and definition of $\mathcal{W}$, we have a constant $\tilde{C}_n$ such that*

$$\|[\mathcal{W}, L_\tau]\| \leq \tilde{C}_n \left(\left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) \|D\tau\|_\infty + \|D^2\tau\|_\infty \right).$$

Theorem 31. *Assume ψ and its first and second order derivatives have decay in $O((1+|x|)^{-n-2})$ and $\int_{\mathbb{R}^n} \psi(x)\,dx = 0$. Then for every $\tau \in \mathbf{C}^2(\mathbb{R}^n)$ with $\|D\tau\|_\infty \leq \frac{1}{2n}$, there exists $C_{m,n} > 0$ such that*

$$\|S^m_{cont,2} f - S^m_{cont,2} L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \leq C_{m,n} \left(\|D\tau\|_\infty + \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^2 \right) \|f\|_2^2.$$

Proof. Like with the previous proof, we have the following bound for some C_m:

$$\begin{aligned}\|S^m_{\text{cont},2} f - S^m_{\text{cont},2} L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} &\leq \|A_2 M \mathcal{W} V_{m-1} f - A_2 M L_\tau \mathcal{W} V_{m-1} f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} + \|A_2 M [\mathcal{W} V_{m-1}, L_\tau] f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} \\ &\leq \|A_2 M \mathcal{W} V_{m-1} f - A_2 M L_\tau \mathcal{W} V_{m-1} f\|_{\mathbf{L}^1(\mathbb{R}^m_+)} + C_m^2 \|[\mathcal{W}, L_\tau]\|^2 \|f\|_2^2.\end{aligned}$$

For the first term, we can mimic the dilation argument to get

$$|A_2 M \mathcal{W} V_{m-1} f - A_2 M L_\tau \mathcal{W} V_{m-1} f| = \left| \|g\|_2^2 - \|L_\tau g\|_2^2 \right|.$$

The difference is the term with the diffeomorphism. Let $y = \gamma(x) = x - \tau(x)$. Then it follows that $\gamma^{-1}(y) = x$ and change of variables implies that

$$\|L_\tau f\|_2^2 = \int_{\mathbb{R}^n} |f(x - \tau(x))|^2\,dx = \int_{\mathbb{R}^n} |f(y)|^2 \frac{dy}{|\det(I - D\tau(\gamma^{-1}(y)))|}.$$

We also have

$$1 - n\|D\tau\|_\infty \leq |\det(I - D\tau(\gamma^{-1}(y)))| \leq 1 + n\|D\tau\|_\infty.$$

Thus, we obtain

$$\begin{aligned}\frac{1}{1+n\|D\tau\|_\infty} \int_{\mathbb{R}^n} |f(y)|^2\,dy \leq \|L_\tau f\|_2^2 \leq \frac{1}{1-n\|D\tau\|_\infty} \int_{\mathbb{R}^n} |f(y)|^2\,dy, \\ \frac{1}{1+n\|D\tau\|_\infty} \|f\|_2^2 \leq \|L_\tau f\|_2^2 \leq \frac{1}{1-n\|D\tau\|_\infty} \|f\|_2^2.\end{aligned}$$

Since we have a bound on $\|D\tau\|_\infty$, we see that

$$\frac{1}{1+n\|D\tau\|_\infty} = \frac{1-n\|D\tau\|_\infty}{1-n^2\|D\tau\|_\infty^2} \geq 1 - n\|D\tau\|_\infty$$

since $1 > 1 - n^2\|D\tau\|_\infty^2 > 0$. Similarly,

$$\frac{1}{1-n\|D\tau\|_\infty} = \frac{1+2n\|D\tau\|_\infty}{1+n\|D\tau\|_\infty - 2n^2\|D\tau\|_\infty^2}.$$

Then

$$1 + n\|D\tau\|_\infty - 2n^2\|D\tau\|_\infty^2 \geq 1 + n\|D\tau\|_\infty - \frac{2n^2}{2n}\|D\tau\|_\infty = 1$$

since $\|D\tau\|_\infty \leq \frac{1}{2n}$. It follows that

$$\frac{1}{1-n\|D\tau\|_\infty} \leq 1 + 2n\|D\tau\|_\infty.$$

Use the lower bound on $\|L_\tau f\|_2^2$ to get

$$\|f\|_2^2 - \|L_\tau f\|_2^2 \le n\|D\tau\|_\infty \|f\|_2^2.$$

The upper bound yields

$$\|L_\tau f\|_2^2 - \|f\|_2^2 \le 2n\|D\tau\|_\infty \|f\|_2^2,$$

and we get

$$\left| \|f\|_2^2 - \|L_\tau f\|_2^2 \right| \le 2n\|D\tau\|_\infty \|f\|_2^2$$

for any $f \in \mathbf{L}^2(\mathbb{R}^n)$.

Now we mimic the argument given for dilation stability to get

$$\|A_2 M \mathcal{W} V_{m-1} f - A_2 M L_\tau \mathcal{W} V_{m-1} f\|_{\mathbf{L}^1(\mathbb{R}_+^m)} \le C\|D\tau\|_\infty \|f\|_2^2$$

for some constant C. For the second term, we use Corollary 30 to get

$$C_m^2 \|[\mathcal{W}, L_\tau]\|^2 \|f\|^2 \le C' \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^2 \|f\|_2^2$$

for some constant C'. We now choose $C_{n,m} = \max\{C', C\}$ to get the desired bound. □

Remark 6. The statement of this theorem looks very similar to a subsection of Theorem 2.12 in [5], which states that there is a constant C such that

$$\|U[\mathcal{P}_{J,m}]f - U[\mathcal{P}_{J,m}]L_\tau f\| \le CK(\tau)(m+1)\|f\|_2^2,$$

with

$$K(\tau) = 2^{-J}\|\tau\|_\infty + \|\nabla\tau\|_\infty \max\left\{ \log \frac{\|\Delta\tau\|_\infty}{\|\nabla\tau\|_\infty}, 1 \right\} + \|\nabla^2\tau\|_\infty.$$

Assuming that we have already proved a commutator bound, one intermediate step is

$$\|U[\mathcal{P}_{J,m}]f - U[\mathcal{P}_{J,m}]L_\tau f\| \le CK(\tau) \sum_{k=1}^{m} \|U[\mathcal{P}_{J,m}]f\|.$$

To highlight the difference in the transforms, the author's proof after the commutator bound is:

$$\sum_{k=1}^{m} \|U[\mathcal{P}_{J,m}]f\| \le (m+1)\|f\|_2^2.$$

This proof is relatively straightforward because $\|U[\mathcal{P}_{J,m}]f\|$ only contains high frequency information (since $j < J$ in their normalization) and $\|U[\mathcal{P}_{J,m}]f\| \le \|U[\mathcal{P}_{J,m-1}]f\|$ since the wavelet transform is nonexpansive. Our bound is less trivial because the corresponding m-layer term for the dyadic transform would be $\|S^m_{\text{dyad},2} f\|_{\mathbf{L}^2(\mathbb{Z}^m)}$, which does not use the averaging operator A_J to contain all frequency information for $j < J$. Nonetheless, the proof for the commutator bound is very similar because the proof in [5] is indexed over $j \in \mathbb{Z}$.

Corollary 32. *Assume ψ and its first and second order derivatives have decay in $O((1+|x|)^{-n-2})$ and $\int_{\mathbb{R}^n} \psi(x)\,dx = 0$. Then for every $\tau \in \mathbf{C}^2(\mathbb{R}^n)$ with $\|D\tau\|_\infty \le \frac{1}{2n}$, there exists $\hat{C}_{m,n} > 0$ such that in the dyadic case we have:*

$$\|S^m_{dyad,2} f - S^m_{dyad,2} L_\tau f\|_{\ell^1(\mathbb{Z}^m)} \le \hat{C}_{m,n} \left(\|D\tau\|_\infty + \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^2 \right) \|f\|_2^2.$$

5.2. Stability to Diffeomorphisms When $1 < q < 2$

Lemma 33. *Let $\gamma(z) = z - \tau(z)$ and define $g(z) = f(\gamma(z))$. Also, define*

$$K_\lambda(x,z) = \det(D\gamma(z))\psi_\lambda(\gamma(x) - \gamma(z)) - \psi_\lambda(x-z),$$

where $\psi_\lambda(t) = \lambda^{-n/2}\psi(t/\lambda)$. If

$$T_\lambda g(x) = \int_{\mathbb{R}^n} g(z)K_\lambda(x,z)\,dz$$

and consider $Tg : \mathbb{R}^n \to \mathbf{L}^2(\mathbb{R}_+, \frac{d\lambda}{\lambda^{n+1}})$ defined by $Tg(x) = (T_\lambda g(x))_{\lambda\in\mathbb{R}_+}$. Then for the Banach space $\mathcal{X} = \mathbf{L}^2(\mathbb{R}_+, \frac{d\lambda}{\lambda^{n+1}})$,

$$\|Tg\|^2_{L^2_{\mathcal{X}}(\mathbb{R}^n)} \leq C_{n,m}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1\right) + \|D^2\tau\|_\infty\right)\|f\|_2^2$$

for some constant $C_{n,m} > 0$.

Proof. Notice that

$$\begin{aligned}
\|Tg\|^2_{L^2_{\mathbb{X}}(\mathbb{R}^n)} &= \int_{\mathbb{R}^n}\int_0^\infty |T_\lambda g(x)|^2\frac{d\lambda}{\lambda^{n+1}}dx \\
&= \int_{\mathbb{R}^n}\int_0^\infty \left|\int_{\mathbb{R}^n} K_\lambda(x,z)g(z)\,dz\right|^2\frac{d\lambda}{\lambda^{n+1}}dx \\
&= \int_{\mathbb{R}^n}\int_0^\infty \left|\int_{\mathbb{R}^n} f(\gamma(z))[\det(D\gamma(z))\psi_\lambda(\gamma(x)-\gamma(z)) - \psi_\lambda(x-z)]\,dz\right|^2\frac{d\lambda}{\lambda^{n+1}}dx \\
&= \int_{\mathbb{R}^n}\int_0^\infty \left|\int_{\mathbb{R}^n} \det(D\gamma(z))f(\gamma(z))\psi_\lambda(\gamma(x)-\gamma(z))\,dz - \int_{\mathbb{R}^n} f(\gamma(z))\psi_\lambda(x-z)\,dz\right|^2\frac{d\lambda}{\lambda^{n+1}}dx.
\end{aligned}$$

Using the change of variables $u = \gamma(z)$, we get

$$\begin{aligned}
\|Tg\|^2_{L^2_{\mathbb{X}}(\mathbb{R}^n)} &= \int_{\mathbb{R}^n}\int_0^\infty |L_\tau(f*\psi_\lambda)(x) - (L_\tau f*\psi_\lambda)(x)|^2\,\frac{d\lambda}{\lambda^{n+1}}\,dx \\
&= \int_{\mathbb{R}^n}\int_0^\infty |[\mathcal{W}_\lambda, L_\tau]f(x)|^2\,\frac{d\lambda}{\lambda^{n+1}}\,dx \\
&= \int_0^\infty\int_{\mathbb{R}^n} |[\mathcal{W}_\lambda, L_\tau]f(x)|^2\,dx\,\frac{d\lambda}{\lambda^{n+1}} \\
&= \int_0^\infty \|[\mathcal{W}_\lambda, L_\tau]f\|_2^2\,\frac{d\lambda}{\lambda^{n+1}} \\
&= \|[\mathcal{W}, L_\tau]f\|_2^2 \\
&\leq C_{n,m}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1\right) + \|D^2\tau\|_\infty\right)\|f\|_2^2,
\end{aligned}$$

where the last inequality follows from the $q = 2$ case. □

Lemma 34. *(Marcinkiewicz Interploation) [19] Let $\mathcal{A}$ and $\mathcal{B}$ be Banach spaces and let $T : \mathcal{A} \to \mathcal{B}$ be a sublinear operator defined on $\mathbf{L}^{p_0}_{\mathcal{A}}(\mathbb{R}^n)$ and $\mathbf{L}^{p_1}_{\mathcal{A}}(\mathbb{R}^n)$ with $0 < p_0 < p_1$. Furthermore, if T satisfies*

$$\|Tf\|_{\mathbf{L}^{p_i,\infty}_{\mathcal{B}}(\mathbb{R}^n)} \leq M_i\|f\|_{\mathbf{L}^{p_i}_{\mathcal{A}}(\mathbb{R}^n)}$$

for $i = 0, 1$, then for all $p \in (p_0, p_1)$,

$$\|Tf\|_{\mathbf{L}^{p}_{\mathcal{B}}(\mathbb{R}^n)} \leq N_p\|f\|_{\mathbf{L}^{p}_{\mathcal{A}}(\mathbb{R}^n)},$$

where N_p only depends on M_0, M_1, and p.

Remark 7. Like with the scalar valued estimate, it can be shown that $N_p = \eta M_0^{\delta} M_1^{1-\delta}$, where

$$\delta = \begin{cases} \dfrac{p_0(p_1 - p)}{p(p_1 - p_0)} & p_1 < \infty, \\ \dfrac{p_0}{p} & p_1 = \infty \end{cases}$$

and

$$\eta = \begin{cases} 2\left(\dfrac{p(p_1 - p_0)}{(p - p_0)(p_1 - p)}\right)^{1/p} & p_1 < \infty, \\ 2\left(\dfrac{p_0}{p - p_0}\right)^{1/p} & p_1 = \infty. \end{cases}$$

Lemma 35. *Let T be the operator defined in Lemma 33. Let $q \in (1,2)$ and $r \in (1,q)$. Then T satisfies*

$$\|Tg\|_{\mathbf{L}^{r,\infty}_{\mathcal{X}}(\mathbb{R}^n)} \leq M_r \|f\|_{\mathbf{L}^r(\mathbb{R}^n)}$$

for some constant $M_r > 0$, which is independent of $\|D\tau\|_\infty$ and $\|D^2\tau\|_\infty$. Furthermore, T also satisfies

$$\|Tg\|^2_{\mathbf{L}^{2,\infty}_{\mathcal{X}}(\mathbb{R}^n)} \leq \tilde{C}_n \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1\right) + \|D^2\tau\|_\infty\right)^2 \|f\|^2_{\mathbf{L}^2(\mathbb{R}^n)}$$

for some constant $\tilde{C}_n > 0$.

Proof. The second inequality obviously follows from strong boundedness of the operator, so we will omit the proof. For the first inequality, the norm satisfies

$$\begin{aligned}
\|Tg(x)\|^2_{\mathcal{X}} &= \int_0^\infty \left| \int_{\mathbb{R}^n} \det(D\gamma(z)) f(\gamma(z)) \psi_\lambda(\gamma(x) - \gamma(z))\, dz - \int_{\mathbb{R}^n} f(\gamma(z)) \psi_\lambda(x - z)\, dz \right|^2 \frac{d\lambda}{\lambda^{n+1}} \\
&= \int_0^\infty \left| \int_{\mathbb{R}^n} f(z) \psi_\lambda(\gamma(x) - z)\, dz - \int_{\mathbb{R}^n} f(\gamma(z)) \psi_\lambda(x - z)\, dz \right|^2 \frac{d\lambda}{\lambda^{n+1}} \\
&\leq 4 \int_0^\infty \left| \int_{\mathbb{R}^n} f(z) \psi_\lambda(\gamma(x) - z)\, dz \right|^2 \frac{d\lambda}{\lambda^2} + 4 \int_0^\infty \left| \int_{\mathbb{R}^n} f(\gamma(z)) \psi_\lambda(x - z)\, dz \right|^2 \frac{d\lambda}{\lambda^{n+1}} \\
&= 4|(Gf)(\gamma(x))|^2 + 4|GL_\tau f(x)|^2.
\end{aligned}$$

It now follows that

$$\|Tg(x)\|_{\mathcal{X}} \leq \sqrt{4|(Gf)(\gamma(x))|^2 + 4|GL_\tau f(x)|^2} \leq 2|(Gf)(\gamma(x))| + 2|GL_\tau f(x)|.$$

For $\delta > 0$, we have

$$\begin{aligned}
m\{\|Tg(x)\|_{\mathcal{X}} > \delta\} &\leq m\{2|(Gf)(\gamma(x))| + 2|GL_\tau f(x)| > \delta\} \\
&\leq \frac{2^r \|(Gf)(\gamma(\cdot))\|^r_{\mathbf{L}^r(\mathbb{R}^n)} + 2^r \|GL_\tau f\|^r_{\mathbf{L}^r(\mathbb{R}^n)}}{\delta} \\
&\leq \frac{2^r}{\delta^r} (\|(Gf)(\gamma(\cdot))\|^r_{\mathbf{L}^r(\mathbb{R}^n)} + \|GL_\tau f\|^r_{\mathbf{L}^r(\mathbb{R}^n)}).
\end{aligned}$$

Since γ is a diffeomorphism, we can use change of variables to get

$$\begin{aligned}
\|(Gf)(\gamma(\cdot))\|^r_{\mathbf{L}^r(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} |Gf(\gamma(x))|^r\, dx \\
&= \int_{\mathbb{R}^n} |Gf(u)|^r \frac{du}{\det[(D\gamma)(\gamma^{-1}(u))]} \\
&\leq 2 \int_{\mathbb{R}^n} |Gf(x)|^r\, dx \\
&= 2\|Gf\|^r_{\mathbf{L}^r(\mathbb{R}^n)}.
\end{aligned}$$

By Theorem 1, we get

$$\|GL_\tau f\|^r_{\mathbf{L}^r(\mathbb{R}^n)} \leq C\|L_\tau f\|^r_{\mathbf{L}^r(\mathbb{R}^n)} \leq 2C\|f\|^r_{\mathbf{L}^r(\mathbb{R}^n)}$$

for some constant C_r dependent on r. Thus, we have

$$m\{\|Tg(x)\|_{\mathcal{X}} > \delta\}^{1/r} \leq \frac{M_r}{\delta}\|f\|_{\mathbf{L}^r(\mathbb{R}^n)}$$

for some constant $M_r > 0$. □

Lemma 36. *Fix $r = \frac{1+q}{2}$ so that $r \in (1, q)$. For some constant $C_{n,q} > 0$, the operator T defined in Lemma 33 satisfies the estimate*

$$\|Tg\|^q_{\mathbf{L}^q_{\mathcal{X}}(\mathbb{R}^n)} \leq C_{n,q}\eta^q M_r^{q\delta}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1\right) + \|D^2\tau\|_\infty\right)^{q(1-\delta)} \|f\|^q_{\mathbf{L}^q(\mathbb{R}^n)},$$

where η and δ come from interpolation, and M_r comes from the constant for weak boundedness in Lemma 35.

Proof. Using the $\mathbf{L}^r(\mathbb{R}^n)$ and $\mathbf{L}^2(\mathbb{R}^n)$ estimates from the previous Lemma, we interpolate using Marcinkiewicz since $\|g\|_r \leq 2\|f\|_r \leq 4\|g\|_r$. In our particular case, since r is fixed, so is δ. It is easy to confirm that $\delta = \frac{1}{1+q} \in \left(\frac{1}{3}, \frac{1}{2}\right)$ when using Marcinkiewicz. □

Theorem 37. *Let $1 < q < 2$ and $pq = 2$. Assume ψ and its first and second order derivatives have decay in $O((1+|x|)^{-n-2})$, and $\int_{\mathbb{R}^n}\psi(x)\,dx = 0$. Then for every $\tau \in \mathbf{C}^2(\mathbb{R}^n)$ with $\|D\tau\|_\infty \leq \frac{1}{2n}$, there exists $C_{n,q} > 0$ such that*

$$\|S_{cont,q}f - S_{cont,q}L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^+)} \leq C_{n,q}\left[\|D\tau\|_\infty + \eta^q M_r^{q\delta}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1\right) + \|D^2\tau\|_\infty\right)^{q(1-\delta)}\right]\|f\|^q_q.$$

Proof. We use the same notation as Proposition 16. Using a nearly identical argument, we get

$$\begin{aligned}
\|S_{\text{cont},q}f - S_{\text{cont},q}L_\tau f\|_{\mathbf{L}^p(\mathbb{R}_+)} &= \|A_qM\mathcal{W}f - A_qM\mathcal{W}L_\tau f\|_{\mathbf{L}^p(\mathbb{R}_+)} \\
&= \|A_qM\mathcal{W}f - A_qML_\tau\mathcal{W}f + A_qML_\tau Wf - A_qM\mathcal{W}L_\tau f\|_{\mathbf{L}^p(\mathbb{R}_+)} \\
&\leq \|A_qM\mathcal{W}f - A_qML_\tau\mathcal{W}f\|_{\mathbf{L}^p(\mathbb{R}_+)} + \|A_qML_\tau\mathcal{W}f - A_qM\mathcal{W}L_\tau f\|_{\mathbf{L}^p(\mathbb{R}_+)} \\
&= \|A_qM\mathcal{W}f - A_qML_\tau\mathcal{W}f\|_{\mathbf{L}^p(\mathbb{R}_+)} + \|A_qML_\tau\mathcal{W}f - A_qM\mathcal{W}L_\tau f\|_{\mathbf{L}^p(\mathbb{R}_+)} \\
&= \|(A_qM - A_qML_\tau)\mathcal{W}f\|_{\mathbf{L}^p(\mathbb{R}_+)} + \|A_qM[\mathcal{W}, L_\tau]f\|_{\mathbf{L}^p(\mathbb{R}_+)}.
\end{aligned}$$

The first term, $\|(A_qM - A_qML_\tau)\mathcal{W}f\|_{\mathbf{L}^p(\mathbb{R}_+)}$, can be bounded using an argument identical to the $q = 2$ case. For the other term,

$$\begin{aligned}
\|A_qM[\mathcal{W}, L_\tau]f\|_{L^p(\mathbb{R}_+)} &= \left(\int_0^\infty |A_qM[\mathcal{W}_\lambda, L_\tau]f|^p \frac{d\lambda}{\lambda^{n+1}}\right)^{1/p} \\
&= \left(\int_0^\infty \|L_\tau f * \psi_\lambda - L_\tau(f * \psi_\lambda)\|_q^{qp} \frac{d\lambda}{\lambda^{n+1}}\right)^{1/p} \\
&= \left(\int_0^\infty \left[\int_{\mathbb{R}^n} |(L_\tau f * \psi_\lambda)(x) - L_\tau(f * \psi_\lambda)(x)|^q\,dx\right]^p \frac{d\lambda}{\lambda^{n+1}}\right)^{1/p}.
\end{aligned}$$

Now, expand convolution and then use change of variables to get

$$
\begin{aligned}
\|A_q M[\mathcal{W}, L_\tau] f\|_{\mathbf{L}^p(\mathbb{R}_+)} &= \left(\int_0^\infty \left[\int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(\gamma(z))(\det(D\gamma(z))\psi_\lambda(\gamma(x) - \gamma(z)) - \psi_\lambda(x - z))\, dz \right|^q dx \right]^p \frac{d\lambda}{\lambda^{n+1}} \right)^{1/p} \\
&= \left(\int_0^\infty \left[\int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} g(z) K_\lambda(x, z)\, dz \right|^q dx \right]^p \frac{d\lambda}{\lambda^{n+1}} \right)^{1/p} \\
&= \left(\int_0^\infty \left[\int_{\mathbb{R}^n} |T_\lambda g(x)|^q\, dx \right]^p \frac{d\lambda}{\lambda^{n+1}} \right)^{1/p} \\
&\leq \int_{\mathbb{R}^n} \left[\int_0^\infty |T_\lambda g(x)|^{qp} \frac{d\lambda}{\lambda^{n+1}} \right]^{1/p} dx \\
&= \int_{\mathbb{R}^n} \left[\int_0^\infty |T_\lambda g(x)|^2 \frac{d\lambda}{\lambda^{n+1}} \right]^{q/2} dx \\
&= \int_{\mathbb{R}^n} \|Tg(x)\|^q_{\mathbf{L}^2\left(\mathbb{R}^+, \frac{d\lambda}{\lambda^{n+1}}\right)} dx \\
&= \|Tg\|^q_{\mathbf{L}^q_{\mathcal{X}}(\mathbb{R}^n)} \\
&\leq C_n \eta^q M_r^{q\delta} \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^{q(1-\delta)} \|f\|_q^q.
\end{aligned}
$$

Thus, the proof is complete. □

Corollary 38. *There exists a constant $C_{n,m} > 0$ such that*

$$
\|S^m_{cont,q} f - S^m_{cont,q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)} \leq C_{n,m} \left[\|D\tau\|_\infty + \eta^q M_r^{q\delta} \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^{q(1-\delta)} \right] \|f\|_q^q.
$$

In the dyadic case, we have a similar corollary as well.

Corollary 39. *There exists a constant $\hat{C}_{n,m} > 0$ such that*

$$
\|S^m_{dyad,q} f - S^m_{dyad,q} L_\tau f\|_{\mathbf{L}^p(\mathbb{Z}^m)} \leq \hat{C}_{n,m} \left[\|D\tau\|_\infty + \eta^q M_r^{q\delta} \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^{q(1-\delta)} \right] \|f\|_q^q.
$$

Remark 8. This bound is not exactly the same as the definition for stability to diffeomorphisms in [5], but the idea is similar since the bound is proportional to $\|D\tau\|_\infty$ and $\|D^2\tau\|_\infty$.

6. Equivariance and Invariance to Rotations

We now consider adding group actions to our scattering transform and prove invariance to rotations. Let $\mathrm{SO}(n)$ be the group of $n \times n$ rotation matrices. Since $\mathrm{SO}(n)$ is a compact Lie group, we can define a Harr measure, say μ, with $\mu(\mathrm{SO}(n)) < \infty$. We say that $f \in \mathbf{L}^2(\mathrm{SO}(n))$ if and only if f is μ-measurable and $\int_{\mathrm{SO}(n)} |f(r)|^2\, d\mu(r) < \infty$.

6.1. Rotation Equivariant Representations

Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a wavelet. Define

$$
\psi_{\lambda,R} = \lambda^{-n/2} \psi(\lambda^{-1} R^{-1} x),
$$

where $R \in \mathrm{SO}(n)$ is a $n \times n$ rotation matrix. The continuous and dyadic wavelet transforms of f are given by

$$\mathcal{W}_{\mathrm{Rot}} f := \{f * \psi_{\lambda,R}(x) : x \in \mathbb{R}^n, \lambda \in (0,\infty), R \in \mathrm{SO}(n)\},$$
$$W_{\mathrm{Rot}} f := \{f * \psi_{j,R}(x) : x \in \mathbb{R}^n, j \in \mathbb{Z}, R \in \mathrm{SO}(n)\}.$$

We will first consider a translation invariant and rotation equivariant formulation of continuous and dyadic one-layer scattering using

$$\mathfrak{S}_{\mathrm{cont},q} f(\lambda, R) := \|f * \psi_{\lambda,R}\|^q_{\mathbf{L}^q(\mathbb{R}^n)},$$
$$\mathfrak{S}_{\mathrm{dyad},q} f(j, R) := \|f * \psi_{j,R}\|^q_{\mathbf{L}^q(\mathbb{R}^n)}.$$

The translation invariance of our representation follows from translation invariance of the norm. For rotation equivariance, notice that if $f_{\tilde{R}}(x) := f(\tilde{R}^{-1}x)$, then we have

$$\mathfrak{S}_{\mathrm{cont},q} f_{\tilde{R}}(\lambda, R) = \mathfrak{S}_{\mathrm{cont},q} f(\lambda, R\tilde{R}),$$
$$\mathfrak{S}_{\mathrm{dyad},q} f_{\tilde{R}}(j, R) = \mathfrak{S}_{\mathrm{dyad},q} f(j, R\tilde{R}).$$

Now suppose we have m layers again. Then we define our m layer transforms by

$$\mathfrak{S}^m_{\mathrm{cont},q} f(\lambda_1, \ldots, \lambda_m, R_1, \ldots, R_m) := \||f * \psi_{\lambda_1,R_1}| * \ldots | * \psi_{\lambda_m,R_m}\|^q_{\mathbf{L}^q(\mathbb{R}^n)},$$
$$\mathfrak{S}^m_{\mathrm{dyad},q} f(j_1, \ldots, j_m, R_1, \ldots, R_m) := \||f * \psi_{j_1,R_1}| * \ldots | * \psi_{j_m,R_m}\|^q_{\mathbf{L}^q(\mathbb{R}^n)}.$$

and rotation equivariance implies

$$\mathfrak{S}^m_{\mathrm{cont},q} f_{\tilde{R}}(\lambda_1, \ldots, \lambda_m, R_1, \ldots, R_m) = \mathfrak{S}^m_{\mathrm{cont},q} f(\lambda_1, \ldots, \lambda_m, R_1\tilde{R}, \ldots, R_m\tilde{R}),$$
$$\mathfrak{S}^m_{\mathrm{dyad},q} f_{\tilde{R}}(j_1, \ldots, j_m, R_1, \ldots, R_m) = \mathfrak{S}^m_{\mathrm{dyad},q} f(j_1, \ldots, j_m, R_1\tilde{R}, \ldots, R_m\tilde{R}).$$

The norm we will use is similar to our previous formulations. Assume $pq = 2$ (this holds when $q = 1$ and $q = 2$ as well). Denote the scattering norm for the continuous transform as

$$\|\mathfrak{S}^m_{\mathrm{cont},q} f\|^p_{\mathbf{L}^p(\mathbb{R}^m_+) \times \mathrm{SO}(n)^m} := \int_0^\infty \int_{\mathrm{SO}(n)} \cdots \int_0^\infty \int_{\mathrm{SO}(n)} \||f * \psi_{j_1,R_1}| * \ldots | * \psi_{j_m,R_m}\|^{qp}_{\mathbf{L}^q(\mathbb{R}^n)} d\mu_1(R_1) \frac{d\lambda_1}{\lambda_1^{n+1}} \ldots d\mu_m(R_n) \frac{d\lambda_m}{\lambda_m^{n+1}}.$$

For the dyadic transform, we denote the norm using

$$\|\mathfrak{S}^m_{\mathrm{dyad},q} f\|^p_{\ell^p(\mathbb{Z}^m) \times \mathrm{SO}(n)^m} := \sum_{j_m \in \mathbb{Z}} \int_{\mathrm{SO}(n)} \cdots \sum_{j_1 \in \mathbb{Z}} \int_{\mathrm{SO}(n)} \||f * \psi_{j_1,R_1}| * \ldots | * \psi_{j_m,R_m}\|^{qp}_{\mathbf{L}^q(\mathbb{R}^n)} d\mu_1(R_1) \ldots d\mu_m(R_n).$$

We'll start by proving that these formulations of the scattering transform are well defined, and prove properties about stability to diffeomorphisms like in previous sections.

Lemma 40. *Let ψ be a wavelet that satisfies properties (24) and (25).*

- *If $1 < q \leq 2$, we have $\mathfrak{S}^m_{cont,q} : \mathbf{L}^q(\mathbb{R}^n) \to \mathbf{L}^p(\mathbb{R}^m_+) \times SO(n)^m$ and $\mathfrak{S}^m_{dyad,q} : \mathbf{L}^2(\mathbb{R}^n) \to \ell^p(\mathbb{Z}^m) \times SO(n)^m$.*
- *If ψ has $n+3$ vanishing moments, then $\mathfrak{S}^m_{cont,1} : \mathbf{L}^1(\mathbb{R}^n) \to \mathbf{L}^2(\mathbb{R}^m_+) \times SO(n)^m$ and $\mathfrak{S}^m_{dyad,1} : \mathbf{L}^1(\mathbb{R}^n) \to \ell^2(\mathbb{Z}^m) \times SO(n)^m$.*
- *If ψ is also a Littlewood-Paley wavelet, we have*

$$\|\mathfrak{S}^m_{cont,2} f\|_{\mathbf{L}^1(\mathbb{R}^m_+) \times SO(n)^m} = \mu(SO(n))^m C^m_\psi \|f\|_2^2,$$
$$\|\mathfrak{S}^m_{dyad,q} f\|_{\ell^1(\mathbb{Z}^m) \times SO(n)^m} = \mu(SO(n))^m \hat{C}^m_\psi \|f\|_2^2.$$

Proof. We prove the first and third claim. The second claim is almost identical to the first claim, so the proof will be omitted for brevity. Note that we will only provide arguments for the continuous scattering transform since the proofs for the dyadic transform are very similar.

By Fubini and boundedness of the m-layer scattering transform, there exists a constant $C_q > 0$, which is dependent on q, such that

$$\begin{aligned}
&\|\mathfrak{S}^m_{\text{cont},q} f\|_{\mathbf{L}^p(\mathbb{R}^m_+)\times \mathrm{SO}(n)^m} \\
&= \int_0^\infty \int_{\mathrm{SO}(n)} \cdots \int_0^\infty \int_{\mathrm{SO}(n)} \||f * \psi_{\lambda_1,R_1}| * \ldots | * \psi_{\lambda_m,R_m}\|^{qp}_{\mathbf{L}^q(\mathbb{R}^n)} d\mu(R_m) \frac{d\lambda_m}{\lambda_m^{n+1}} \cdots d\mu(R_1) \frac{d\lambda_1}{\lambda_1^{n+1}} \\
&= \int_{\mathrm{SO}(n)} \cdots \int_{\mathrm{SO}(n)} \left(\int_0^\infty \cdots \int_0^\infty \||f * \psi_{\lambda_1,R_1}| * \ldots | * \psi_{\lambda_m,R_m}\|^{qp}_{\mathbf{L}^q(\mathbb{R}^n)} \frac{d\lambda_m}{\lambda_m^{n+1}} \cdots \frac{d\lambda_1}{\lambda_1^{n+1}} \right) d\mu(R_1) \cdots d\mu(R_n) \\
&\leq \int_{\mathrm{SO}(n)} \cdots \int_{\mathrm{SO}(n)} C_q^m \|f\|_q^{qp}\, d\mu(R_1) \cdots d\mu(R_m) \\
&= C_q^m \mu(\mathrm{SO}(n))^m \|f\|_q^q
\end{aligned}$$

because each ψ_{λ_i,R_i} is still a wavelet with sufficient decay even if the rotation is applied.

For the third claim, we see that

$$\begin{aligned}
&\|\mathfrak{S}^m_{\text{cont},2} f\|_{\mathbf{L}^1(\mathbb{R}^m_+)\times \mathrm{SO}(n)^m} \\
&= \int_{\mathrm{SO}(n)} \cdots \int_{\mathrm{SO}(n)} \left(\int_0^\infty \cdots \int_0^\infty \||f * \psi_{\lambda_1,R_1}| * \ldots | * \psi_{\lambda_m,R_m}\|^{2}_{\mathbf{L}^2(\mathbb{R}^n)} \frac{d\lambda_m}{\lambda_m^{n+1}} \cdots \frac{d\lambda_1}{\lambda_1^{n+1}} \right) d\mu(R_1) \cdots d\mu(R_n) \\
&= \int_{\mathrm{SO}(n)} \cdots \int_{\mathrm{SO}(n)} C_\psi^m \|f\|_2^2\, d\mu(R_1) \cdots d\mu(R_m) \\
&= \mu(\mathrm{SO}(n))^m C_\psi^m \|f\|_2^2.
\end{aligned}$$

□

Theorem 41. *Let $\tau(x) = cx$ and let $L_\tau f(x) = f((1-c)x)$. Suppose that ψ is a wavelet that satisfies the conditions of Lemma 11. Then there exist constants $\tilde{K}_{n,m,q}$ and $\tilde{K}'_{n,m,q}$ dependent only on n, m, and q such that*

$$\|\mathfrak{S}^m_{cont,q} f - \mathfrak{S}^m_{cont,q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)\times SO(n)^m} \leq c \cdot \tilde{K}_{n,m,q} \|f\|_q^q$$

and

$$\|\mathfrak{S}^m_{dyad,q} f - \mathfrak{S}^m_{dyad,q} L_\tau f\|_{\ell^p(\mathbb{Z}^m)\times SO(n)^m} \leq c \cdot \tilde{K}'_{n,m,q} \|f\|_q^q$$

for any $c < \frac{1}{2n}$. Additionally, there exists $\tilde{H}_{m,n}$ and $\tilde{H}'_{m,n}$ such that

$$\|\mathfrak{S}^m_{cont,1} f - \mathfrak{S}^m_{cont,1} L_\tau f\|_{\mathbf{L}^2(\mathbb{R}^m_+)\times SO(n)^m} \leq c \cdot \tilde{H}_{m,n} \|f\|_{\mathbb{H}^1(\mathbb{R}^n)}$$

and

$$\|\mathfrak{S}^m_{dyad,1} f - \mathfrak{S}^m_{dyad,1} L_\tau f\|_{\ell^2(\mathbb{Z}^m)\times SO(n)^m} \leq c \cdot \tilde{H}'_{m,n} \|f\|_{\mathbb{H}^1(\mathbb{R}^n)}$$

Proof. We provide the proof when $q \neq 2$ in the continuous case. We can use Fubini and Corollary 26 to get

$$\begin{aligned}
&\|\mathfrak{S}^m_{\text{cont},q} f - \mathfrak{S}^m_{\text{cont},q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)\times \mathrm{SO}(n)^m} \\
&= \int_{\mathrm{SO}(n)} \cdots \int_{\mathrm{SO}(n)} \left(\int_0^\infty \cdots \int_0^\infty |\mathfrak{S}^m_{\text{cont},q} f - \mathfrak{S}^m_{\text{cont},q} L_\tau f| \frac{d\lambda_m}{\lambda_m^{n+1}} \cdots \frac{d\lambda_1}{\lambda_1^{n+1}} \right) d\mu(R_1) \cdots d\mu(R_n) \\
&\leq \int_{\mathrm{SO}(n)} \cdots \int_{\mathrm{SO}(n)} C_q K_{n,m,q} \|f\|_q^q\, d\mu(R_1) \cdots d\mu(R_n) \\
&= c \cdot C_q \cdot \mathrm{SO}(n)^m K_{n,m,q} \|f\|_q^q.
\end{aligned}$$

□

Theorem 42. *Let $\tau \in C^2(\mathbb{R}^n)$ and let $L_\tau f(x) = f(x - \tau(x))$. Suppose that ψ is a wavelet such that the wavelet and all its first and second partial derivatives have $O((1+|x|)^{-n-2})$ decay. When $q \in (1,2)$, there exists a constant $C_{n,m,q}$ dependent on $\mu(SO(n))$, n, m, and q such that*

$$\|\mathfrak{S}^m_{cont,q} f - \mathfrak{S}^m_{cont,q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)\times SO(n)^m} \leq C_{n,m,q}\left[\|D\tau\|_\infty + \eta^q M_r^{q\delta}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right) + \|D^2\tau\|_\infty\right)^{q(1-\delta)}\right]\|f\|_q^q.$$

Also, when $q = 2$, there exists a constant $C_{n,m}$ such that

$$\|\mathfrak{S}^m_{cont,2} f - \mathfrak{S}^m_{cont,2} L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)\times SO(n)^m} \leq C_{n,m}\left[\|D\tau\|_\infty + \left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right) + \|D^2\tau\|_\infty\right)^2\right]\|f\|_2^2.$$

Proof. We only provide the proof when $q \neq 2$. The idea is the same as before:

$$\begin{aligned}
&\|\mathfrak{S}^m_{\text{cont},q} f - \mathfrak{S}^m_{\text{cont},q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)\times \mathrm{SO}(n)^m} \\
&= \int_{\mathrm{SO}(n)}\cdots\int_{\mathrm{SO}(n)}\left(\int_0^\infty\cdots\int_0^\infty |\mathfrak{S}^m_{\text{cont},q} f - \mathfrak{S}^m_{\text{cont},q} L_\tau f|\frac{d\lambda_m}{\lambda_m^{n+1}}\cdots\frac{d\lambda_1}{\lambda_1^{n+1}}\right)d\mu(R_1)\cdots d\mu(R_n) \\
&\leq \int_{\mathrm{SO}(n)}\cdots\int_{\mathrm{SO}(n)} C_{n,m,q}\left[\|D\tau\|_\infty + \eta^q M_r^{q\delta}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right) + \|D^2\tau\|_\infty\right)^{q(1-\delta)}\right]\|f\|_q^q\, d\mu(R_1)\cdots d\mu(R_n) \\
&= C_{n,m,q}\left[\|D\tau\|_\infty + \eta^q M_r^{q\delta}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right) + \|D^2\tau\|_\infty\right)^{q(1-\delta)}\right]\|f\|_q^q.
\end{aligned}$$

□

Corollary 43. *Let $\tau \in C^2(\mathbb{R}^n)$ and let $L_\tau f(x) = f(x - \tau(x))$. Suppose that ψ is a wavelet such that the wavelet and all its first and second partial derivatives have $O((1+|x|)^{-n-2})$ decay. When $q \in (1,2)$, there exists a constant $\tilde{C}_{n,p}$ dependent on $\mu(SO(n))$, n, m, and q such that*

$$\|\mathfrak{S}^m_{dyad,q} f - \mathfrak{S}^m_{dyad,q} L_\tau f\|_{\ell^p(\mathbb{Z}^m)\times SO(n)^m} \leq \tilde{C}_{n,m,q}\left[\|D\tau\|_\infty + \eta^q M_r^{q\delta}\left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right) + \|D^2\tau\|_\infty\right)^{q(1-\delta)}\right]\|f\|_q^q.$$

Also, when $q = 2$, there exists a constant $\tilde{C}_{n,m}$ such that

$$\|\mathfrak{S}^m_{dyad,2} f - \mathfrak{S}^m_{dyad,2} L_\tau f\|_{\mathbf{L}^1(\mathbb{Z}^m)\times SO(n)^m} \leq \tilde{C}_{n,m}\left[\|D\tau\|_\infty + \left(\|D\tau\|_\infty\left(\log\frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty}\vee 1\right) + \|D^2\tau\|_\infty\right)^2\right]\|f\|_2^2.$$

6.2. Rotation Invariant Representations

The representation before was rotation equivariant, but in some tasks, we would also like rotation invariance. In [5], the authors choose to integrate over each group action in a group of transformations. However, this will remove the information the relative angles between each action if we have multiple layers in our transform.

In the case of one layer, since there is only one angle, we use a similar formulation to [5] and define continuous and dyadic scattering transforms for rotation invariance as

$$\mathscr{S}_{\text{cont},q} f(\lambda) = \int_{\mathrm{SO}(n)} \|f * \psi_{\lambda,R}\|^q_{\mathbf{L}^q(\mathbb{R}^n)}\mu(R),$$
$$\mathscr{S}_{\text{dyad},q} f(j) = \int_{\mathrm{SO}(n)} \|f * \psi_{j,R}\|^q_{\mathbf{L}^q(\mathbb{R}^n)}\mu(R).$$

The corresponding norms are given by

$$\|\mathscr{S}_{\mathrm{cont},q} f\|^p_{\mathbf{L}^p(\mathbb{R}_+)} := \int_0^\infty \left[\int_{\mathrm{SO}(n)} \|f * \psi_{\lambda,R}\|^q_{\mathbf{L}^q(\mathbb{R}^n)} \mu(R)\right]^p \frac{d\lambda}{\lambda^{n+1}},$$
$$\|\mathscr{S}_{\mathrm{dyad},q} f\|^p_{\ell^p(\mathbb{Z})} := \sum_{j\in\mathbb{Z}} \left[\int_{\mathrm{SO}(n)} \|f * \psi_{j,R}\|^q_{\mathbf{L}^q(\mathbb{R}^n)} \mu(R)\right]^p.$$

Now we generalize to the case where $m \geq 2$. Let $R_1, \dots, R_m \in \mathrm{SO}(n)$. Define

$$\mathscr{S}^m_{\mathrm{cont},q} f(\lambda_1, \dots, \lambda_m, R_2, \dots, R_m) := \int_{\mathrm{SO}(n)} \| |f * \psi_{\lambda_1,R_2R_1}| * \dots * |\psi_{\lambda_m,R_mR_1}\|^q_{\mathbf{L}^q(\mathbb{R}^n)} \, \mu(R_1),$$
$$\mathscr{S}^m_{\mathrm{dyad},q} f(j_1, \dots, j_m, R_2, \dots, R_m) := \int_{\mathrm{SO}(n)} \| |f * \psi_{j_1,R_2R_1}| * \dots | * \psi_{j_m,R_mR_1}\|^q_{\mathbf{L}^q(\mathbb{R}^n)} \, \mu(R_1).$$

The norm for the continuous transform the norm $\|\mathscr{S}^m_{\mathrm{cont},q} f\|^p_{\mathbf{L}^p(\mathbb{R}^m_+)\times \mathrm{SO}(n)^m}$ is given by

$$\int_0^\infty \int_{\mathrm{SO}(n)} \cdots \int_0^\infty \int_{\mathrm{SO}(n)} \left(\int_0^\infty \left[\mathscr{S}^m_{\mathrm{cont},q} f(\lambda_1, \dots, \lambda_m, R_1, \dots, R_m)\right]^p \frac{d\lambda_1}{\lambda_1^{n+1}}\right) d\mu_2(R_2) \frac{d\lambda_2}{\lambda_2^{n+1}} \dots d\mu_m(R_m) \frac{d\lambda_m}{\lambda_m^{n+1}}.$$

For the dyadic transform, the norm $\|\mathscr{S}^m_{\mathrm{dyad},q} f\|^p_{\ell^p(\mathbb{Z})\times \mathrm{SO}(n)^m}$ is given by

$$\sum_{j_m\in\mathbb{Z}} \int_{\mathrm{SO}(n)} \cdots \sum_{j_2\in\mathbb{Z}} \left(\int_{\mathrm{SO}(n)} \sum_{j_1\in\mathbb{Z}} \left[\mathscr{S}^m_{\mathrm{dyad},q} f(\lambda_1, \dots, \lambda_m, R_1, \dots, R_m)\right]^p d\mu_1(R_1)\right) d\mu_2(R_2) \dots d\mu_m(R_,).$$

Like before, we will discuss the well-definedness and stability of these operators to diffeomorphisms. The proofs will be omitted since they follow directly from the previous sections by using Jensen's inequality to take in the p^{th} power and accumulating a power of $\mu(\mathrm{SO}(n))$ in the final constant.

Lemma 44. *Let ψ be a wavelet that satisfies properties (24) and (25).*

- *If $1 < q \leq 2$, we have $\mathscr{S}^m_{cont,q} : \mathbf{L}^q(\mathbb{R}^n) \to \mathbf{L}^p(\mathbb{R}^m_+) \times SO(n)^{m-1}$ and $\mathscr{S}^m_{dyad,q} : \mathbf{L}^2(\mathbb{R}^n) \to \ell^p(\mathbb{Z}^m) \times SO(n)^{m-1}$.*
- *If $q = 1$ and ψ has $n+3$ vanishing moments, then $\mathscr{S}^m_{cont,1} : \mathbf{L}^1(\mathbb{R}^n) \to \mathbf{L}^2(\mathbb{R}^m_+) \times SO(n)^{m-1}$ and $\mathscr{S}^m_{dyad,1} : \mathbf{L}^1(\mathbb{R}^n) \to \ell^2(\mathbb{Z}^m) \times SO(n)^{m-1}$.*
- *If $q = 2$ and ψ is also a littlewood paley wavelet, we have $\|\mathscr{S}^m_{dyad,2} f\|_{\ell^1(\mathbb{Z}^m)\times SO(n)^{m-1}} = \mu(SO(n))^m C^m_\psi \|f\|^2_2$ and $\|\mathscr{S}^m_{cont,2} f\|_{\mathbf{L}^1(\mathbb{R}^m_+)\times SO(n)^{m-1}} = \mu(SO(n))^m \hat{C}^m_\psi \|f\|^2_2$.*

Theorem 45. *Let $\tau(x) = cx$ and let $L_\tau f(x) = f((1-c)x)$. Suppose that ψ is a wavelet that satisfies the conditions of Lemma 11. Then there exist constants $\hat{K}_{n,m,q}$ and $\hat{K}'_{n,m,q}$ dependent only on n, m, and q such that*

$$\|\mathscr{S}^m_{cont,q} f - \mathscr{S}^m_{cont,q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)\times SO(n)^{m-1}} \leq c \cdot \hat{K}_{n,m,q} \|f\|^q_q$$

and

$$\|\mathscr{S}^m_{dyad,q} f - \mathscr{S}^m_{dyad,q} L_\tau f\|_{\ell^p(\mathbb{Z}^m)\times SO(n)^{m-1}} \leq c \cdot \hat{K}'_{n,m,q} \|f\|^q_q$$

for any $c < \frac{1}{2n}$. Additionally, there exists $\hat{H}_{m,n}$ and $\hat{H}'_{m,n}$ such that

$$\|\mathscr{S}^m_{cont,1} f - \mathscr{S}^m_{cont,1} L_\tau f\|_{\mathbf{L}^2(\mathbb{R}^m_+)\times SO(n)^{m-1}} \leq c \cdot \hat{H}_{m,n} \|f\|_{\mathbb{H}^1(\mathbb{R}^n)}$$

and

$$\|\mathscr{S}^m_{dyad,1} f - \mathscr{S}^m_{dyad,1} L_\tau f\|_{\ell^2(\mathbb{Z}^m)\times SO(n)^{m-1}} \leq c \cdot \hat{H}'_{m,n} \|f\|_{\mathbb{H}^1(\mathbb{R}^n)}$$

Theorem 46. *Let $\tau \in C^2(\mathbb{R}^n)$ and define $L_\tau f(x) = f(x - \tau(x))$ with $\|D\tau\|_\infty < \frac{1}{2n}$. Suppose that ψ is a wavelet such that the wavelet and all its first and second partial derivatives have $O((1+|x|)^{-n-2})$ decay. Then there exists a constant $C_{m,n}$ dependent on $\mu(SO(n))$, m, and n such that*

$$\|\mathscr{S}^m_{cont,2} f - \mathscr{S}^m_{cont,2} L_\tau f\|_{\mathbf{L}^1(\mathbb{R}^m_+)\times SO(n)^{m-1}} \leq C_{m,n} \left(\|D\tau\|_\infty + \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^2 \right) \|f\|_2^2.$$

Corollary 47. *Let $\tau \in C^2(\mathbb{R}^n)$ and define $L_\tau f(x) = f(x - \tau(x))$ with $\|D\tau\|_\infty < \frac{1}{2n}$. Suppose that ψ is a wavelet such that the wavelet and all its first and second partial derivatives have $O((1+|x|)^{-n-2})$ decay.. Then there exists a constant $C_{m,n}$ dependent on $\mu(SO(n))$, m, and n such that*

$$\|\mathscr{S}^m_{dyad,2} f - \mathscr{S}^m_{dyad,2} L_\tau f\|_{\ell^1(\mathbb{Z}^m)\times SO(n)^{m-1}} \leq \hat{C}_{m,n} \left(\|D\tau\|_\infty + \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^2 \right) \|f\|_2^2.$$

Corollary 48. *Let $\tau \in C^2(\mathbb{R}^n)$ and define $L_\tau f(x) = f(x - \tau(x))$ with $\|D\tau\|_\infty < \frac{1}{2n}$. Suppose that ψ is a wavelet such that the wavelet and all its first and second partial derivatives have $O((1+|x|)^{-n-2})$ decay. Then there exists a constant $C_{m,n,q}$ dependent on $\mu(SO(n))$, m, n, and q such that*

$$\|\mathscr{S}_{cont,q} f - \mathscr{S}^m_{cont,q} L_\tau f\|_{\mathbf{L}^p(\mathbb{R}^m_+)\times SO(n)^{m-1}} \leq C_{m,n,q} \left[\|D\tau\|_\infty + \eta^q M_r^{q\delta} \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^{q(1-\delta)} \right] \|f\|_q^q.$$

Corollary 49. *Let $\tau \in C^2(\mathbb{R}^n)$ and define $L_\tau f(x) = f(x - \tau(x))$ with $\|D\tau\|_\infty < \frac{1}{2n}$. Suppose that ψ is a wavelet such that the wavelet and all its first and second partial derivatives have $O((1+|x|)^{-n-2})$ decay. Then there exists a constant $\hat{C}_{m,n,q}$ dependent on $\mu(SO(n))$, m, n, and q such that*

$$\|\mathscr{S}^m_{dyad,q} f - \mathscr{S}^m_{dyad,q} L_\tau f\|_{\ell^p(\mathbb{Z}^m)\times SO(n)^{m-1}} \leq \hat{C}_{m,n,q} \left[\|D\tau\|_\infty + \eta^q M_r^{q\delta} \left(\|D\tau\|_\infty \left(\log \frac{\|\Delta\tau\|_\infty}{\|D\tau\|_\infty} \vee 1 \right) + \|D^2\tau\|_\infty \right)^{q(1-\delta)} \right] \|f\|_q^q.$$

7. Conclusion

We have formulated operators that are translation in variant in $\mathbf{L}^q(\mathbb{R}^n)$, proven these operators are Lipschitz continuous to the action of C^2 diffeomorphisms when $1 < q \leq 2$ with respect to certain norms, and used these results to formulate rotation invariant/equivariant operators on $\mathbf{L}^q(\mathbb{R}^n)$ that are Lipschitz continuous to the action of C^2 diffeomorphisms. One question that was left unanswered was if Lipschitz continuity holds for general diffeomorphisms when $q = 1$. This question is harder to answer because $f \in \mathbb{H}^1(\mathbb{R}^n)$ does not necessarily imply that $L_\tau f \in \mathbb{H}^1(\mathbb{R}^n)$. The kernel for the commutator is also singular, which would mean extension theorems for Hardy spaces could not be used. The answer is most likely no, but we did not construct a counterexample.

8. Acknowledgments

We would like to thank Michael Perlmutter for providing suggestions that improved the clarity of our initial draft.

References

[1] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, Communications of the ACM 60 (6) (2017) 84–90.

[2] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: International Conference on Learning Representations, 2015.

[3] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[4] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[5] S. Mallat, Group invariant scattering, Communications on Pure and Applied Mathematics 65 (10) (2012) 1331–1398.
[6] F. Gao, G. Wolf, M. Hirn, Geometric scattering for graph data analysis, in: International Conference on Machine Learning, PMLR, 2019, pp. 2122–2131.
[7] M. Perlmutter, F. Gao, G. Wolf, M. Hirn, Geometric wavelet scattering networks on compact Riemannian manifolds, in: Mathematical and Scientific Machine Learning, PMLR, 2020, pp. 570–604.
[8] M. Hirn, S. Mallat, N. Poilvert, Wavelet scattering regression of quantum chemical energies, Multiscale Modeling & Simulation 15 (2) (2017) 827–863.
[9] M. Eickenberg, G. Exarchakis, M. Hirn, S. Mallat, L. Thiry, Solid harmonic wavelet scattering for predictions of molecule properties, The Journal of chemical physics 148 (24) (2018) 241732.
[10] P. Sinz, M. W. Swift, X. Brumwell, J. Liu, K. J. Kim, Y. Qi, M. Hirn, Wavelet scattering networks for atomistic systems with extrapolation of material properties, The Journal of Chemical Physics 153 (8) (2020) 084109.
[11] F. Gama, A. Ribeiro, J. Bruna, Stability of graph scattering transforms, Advances in Neural Information Processing Systems 32.
[12] F. Nicola, S. I. Trapasso, Stability of the scattering transform for deformations with minimal regularity, arXiv preprint arXiv:2205.11142.
[13] T. Wiatowski, H. Bölcskei, A mathematical theory of deep convolutional neural networks for feature extraction, IEEE Transactions on Information Theory 64 (3) (2017) 1845–1866.
[14] M. Koller, J. Großmann, U. Monich, H. Boche, Deformation stability of deep convolutional neural networks on Sobolev spaces, in: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2018, pp. 6872–6876.
[15] W. Czaja, W. Li, Analysis of time-frequency scattering transforms, Applied and Computational Harmonic Analysis 47 (1) (2019) 149–171.
[16] A. Bietti, J. Mairal, Group invariance, stability to deformations, and complexity of deep convolutional representations, The Journal of Machine Learning Research 20 (1) (2019) 876–924.
[17] L. Grafakos, Modern fourier analysis, Vol. 250, Springer, 2009.
[18] L. Grafakos, Classical fourier analysis, Vol. 249, Springer, 2014.
[19] J. García-Cuerva, J. R. De Francia, Weighted norm inequalities and related topics, Elsevier, 1985.
[20] E. M. Stein, T. S. Murphy, Harmonic Analysis (PMS-43): Real-Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton University Press, 1993.