SIAM J. IMAGING SCIENCES
Vol. 6, No. 2, pp. 984–998
© 2013 Society for Industrial and Applied Mathematics

Decay Properties of Riesz Transforms and Steerable Wavelets*

John Paul Ward†, Kunal Narayan Chaudhury‡, and Michael Unser†

Abstract. The Riesz transform is a natural multidimensional extension of the Hilbert transform, and it has been the object of study for many years due to its nice mathematical properties. More recently, the Riesz transform and its variants have been used to construct complex wavelets and steerable wavelet frames in higher dimensions. The flip side of this approach, however, is that the Riesz transform of a wavelet often has slow decay. One can nevertheless overcome this problem by requiring the original wavelet to have sufficient smoothness, decay, and vanishing moments. In this paper, we derive necessary conditions in terms of these three properties that guarantee the decay of the Riesz transform and its variants, and, as an application, we show how the decay of the popular Simoncelli wavelets can be improved by appropriately modifying their Fourier transforms. By applying the Riesz transform to these new wavelets, we obtain steerable frames with rapid decay.

Key words. Riesz transform, steerable wavelets, singular integrals

AMS subject classifications. 32A55, 42C40

DOI. 10.1137/120864143

1. Introduction.

The Riesz transform of a real-valued function $f(\boldsymbol{x})$ on $\mathbf{R}^d$ is the vector-valued function $\mathcal{R}f(\boldsymbol{x}) = (\mathcal{R}_1 f(\boldsymbol{x}), \ldots, \mathcal{R}_d f(\boldsymbol{x}))$ given by

$$\mathcal{R}_i f(\boldsymbol{x}) = C_d \lim_{\epsilon \to 0} \int_{\|\boldsymbol{x}-\boldsymbol{y}\|>\epsilon} f(\boldsymbol{y}) \frac{x_i - y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}} \mathrm{d}\boldsymbol{y}, \tag{1.1}$$

where C_d is an absolute constant depending on the dimension d [5]. In this formula, we can see that the Riesz transform is defined by a principal value integral, as is the case for the Hilbert transform. Roughly speaking, the ith component of the Riesz transform is obtained through the convolution of $f(\boldsymbol{x})$ with the kernel $x_i/\|\boldsymbol{x}\|^{d+1}$, where $\|\boldsymbol{x}\|$ is the Euclidean norm of $\boldsymbol{x}$, and x_i are its coordinates. The truncation of the integral in (1.1) is used to contain the singularity of the kernel at the origin, by systematically "chopping" it off around the origin. This added detail makes the analysis of the transform quite involved.

Though the transform appears arcane at first sight, it tends to emerge naturally in various physical and mathematical settings due to its fundamental nature. It is essentially the unique scalar-to-vector–valued transform that is unitary (conserves energy) and is simultaneously invariant to the fundamental transformations of translation, scaling, and rotation. For example,

*Received by the editors January 30, 2012; accepted for publication (in revised form) March 27, 2013; published electronically May 21, 2013. This research was funded in part by ERC grant ERC-2010-AdG 267439-FUN-SP and the Swiss National Science Foundation under grant PBELP2-135867.
http://www.siam.org/journals/siims/6-2/86414.html
†Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland (john.ward@epfl.ch, michael.unser@epfl.ch).
‡Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08540 (kchaudhu@math.princeton.edu).

the role of the Riesz transform in the theory of differential equations and multidimensional Fourier analysis has long been recognized in mathematics, where the above-mentioned invariances play a central role; see, e.g., [5, 20, 21]. The invariance of the transform to translations, namely, that $\mathcal{R}[f(\cdot - \boldsymbol{u})](\boldsymbol{x}) = [\mathcal{R}f](\boldsymbol{x} - \boldsymbol{u})$, is immediately seen from (1.1). On the other hand, its invariance to scaling is confirmed by the change of variable $\boldsymbol{x} \mapsto s\boldsymbol{x}$ in (1.1), where $s > 0$. The other properties are, however, not obvious at first sight. Instead, these are better seen from its Fourier transform, which is given by (see, e.g., [20])

$$\widehat{\mathcal{R}_i f}(\boldsymbol{\omega}) = -j\frac{\omega_i}{\|\boldsymbol{\omega}\|}\widehat{f}(\boldsymbol{\omega}). \tag{1.2}$$

From this and Parseval's theorem, we see the sum of the squared $\mathbf{L}^2$-norms of $\mathcal{R}_i f(\boldsymbol{x})$ equals the squared $\mathbf{L}^2$-norm of $f(\boldsymbol{x})$. That is, the energy of $f(\boldsymbol{x})$ is divided between the different Riesz components. The Fourier expression also tells us that $f(\boldsymbol{x})$ can be recovered from its Riesz transform using a simple reconstruction, namely, as

$$f(\boldsymbol{x}) = -\sum_{i=1}^{d} \mathcal{R}_i[\mathcal{R}_i f(\boldsymbol{x})].$$

Using the invariance of the Fourier transform with rotations, it can also be verified from (1.2) that the Riesz transform of the rotation of $f(\boldsymbol{x})$ is the (vector-field) rotation of $\mathcal{R}f(\boldsymbol{x})$.

The Riesz transform is a natural multidimensional extension of the Hilbert transform, obtained by setting $d = 1$ in (1.1). The fact that the Hilbert transform plays a central role in determining the (complex) analytic counterpart of real-valued signals, and, particularly, in determining its instantaneous phase and frequency, has long been known in the signal processing and wavelet community; see, e.g., [1, 4, 9, 17]. The introduction of the Riesz transform to the community, however, has been more recent. In their pioneering work [7, 8], Felsberg and Sommer used the transform to define a two-dimensional generalization of the analytic signal, which they called the *monogenic signal*. At around this time, Larkin, Bone, and Oldfield introduced a complexified version of the Riesz transform in two dimensions [14, 15]. They showed that this complex transform could be used for demodulating interferrograms and for analyzing fringe patterns in optics. Felsberg's work on the monogenic signal triggered a series of research works relating to the applications of the Riesz transform in the design of wavelets. For example, this inspired Metikas and Olhede to realize a monogenic version of the continuous wavelet transform in [16]. Further work, concerning filters based on Riesz transforms, can be found in [6, 12, 13].

More recently, in a series of papers by Unser et al. [22, 23, 24], it was shown how certain functional properties of the transform, particularly its invariances to translations, scaling, and rotations, and its connection with derivatives and the Laplacian, could be used for realizing complex (monogenic) wavelet bases and *steerable* wavelet frames in higher dimensions. This provided a natural functional counterpart to Simoncelli's steerable filterbanks [19], which is a popular framework for decomposing images at multiple scales and orientations. The key connection provided by Unser et al. was that the steerable wavelets of Simoncelli could be realized as the Riesz transform of appropriate bandpass functions. This powerful functional interpretation allowed them to enrich the wavelet design along several directions. For example,

the authors proposed the so-called higher-order Riesz transforms obtained through repeated applications of the various components of the Riesz transforms. They showed that these higher-order transforms could be applied to a given wavelet to obtain a larger palette of steerable wavelets. When applied on the data, these frames act as (multiscale) partial derivatives. The important point was that all of this could be done using fast and stable filterbank algorithms.

An aspect of this theory that has not previously been studied in depth is how the Riesz transform affects the decay of a wavelet. The importance of this issue lies in the fact that a slowly decaying wavelet can produce poor approximations that suffer from problems such as ringing artifacts and boundary effects. Here, we propose that good decay of the Riesz transform $\mathcal{R}f(\boldsymbol{x})$ is determined by the smoothness, decay, and vanishing moments of $f(\boldsymbol{x})$. In fact, this observation explains the poor spatial decay of some of the earlier wavelet constructions involving the Riesz transform. For example, Felsberg applied the transform on the Poisson kernel which has no vanishing moments [7, 8], and Simoncelli used it on wavelets with poor spatial decay (due to the nonsmooth cut-off in the spectrum) [19].

In this paper, we perform a quantitative study of the decay property of the Riesz transform and its higher-order variants. We expand upon the work in [3], where the intimate connection between the Hilbert transform (the one-dimensional Riesz transform) and wavelets was investigated. It was observed there that for a wavelet with a large number of vanishing moments and a reasonably good decay, its Hilbert transform automatically has an ultrafast decay. We extend the results in [3] first to the Riesz transform and then to its higher-order counterparts. While the main principles are the same, we are required to use a different approach in the higher dimensional setting. The reason for this is that when $d = 1$, the Riesz multiplier has the particularly simple form $1/\pi x$, which does not involve the term $|x|$. Also, the concept of higher-order transforms is trivial for $d = 1$, since by applying the Hilbert transform twice, we get back the negative of the original wavelet. The higher-order transforms are thus of significance only for $d \geq 2$. The analysis of these transforms presents additional technical challenges due to the asymmetry between the decay of the original wavelet and its Riesz transform. As a consequence, the results for the higher-order transforms cannot simply be obtained by a direct application of the results for a single Riesz transform.

2. Notation. We will follow the notation in [22, 23, 24]. The components of the higher-order Riesz transforms are obtained by iterated applications of the component transforms. Thus, given a multi-index of nonnegative integers $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$, we define the higher-order Riesz transform by

$$\mathcal{R}^{\boldsymbol{\alpha}} f(\boldsymbol{x}) = \sqrt{\frac{|\boldsymbol{\alpha}|!}{\boldsymbol{\alpha}!}} \mathcal{R}_1^{\alpha_1} \cdots \mathcal{R}_d^{\alpha_d} f(\boldsymbol{x}),$$

where $\mathcal{R}_i^k$ denotes the k-fold application of $\mathcal{R}_i$ to $f(\boldsymbol{x})$, and $|\boldsymbol{\alpha}|$ is the order of the transform. Here we are using the notation $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_d$ and $\boldsymbol{\alpha}! = \alpha_1! \cdots \alpha_d!$. Also, we will write the partial derivative $D^{\boldsymbol{\alpha}} f(\boldsymbol{x})$ as

$$D^{\boldsymbol{\alpha}} f(\boldsymbol{x}) = \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}} f(\boldsymbol{x})$$

for a given multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$. The inner-product between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ will be denoted by $\boldsymbol{x} \cdot \boldsymbol{y}$, and $\boldsymbol{x}^{\boldsymbol{\alpha}}$ will represent the monomial $x_1^{\alpha_1} \cdots x_d^{\alpha_d}$.

It is clear that the various components of $\mathcal{R}f(\boldsymbol{x})$ have similar analytic properties. We therefore fix i once and for all and define the Riesz kernel $R(\boldsymbol{y}) = y_i/\|\boldsymbol{y}\|^{d+1}$. Our goal in this paper is to study the Riesz transforms of well-behaved functions, with a certain number of vanishing moments. By well-behaved, we mean that they should be sufficiently smooth and must have enough decay. To this end, we define the class $L_N(\mathbf{R}^d)$ to be the collection of differentiable functions $f(\boldsymbol{x})$ such that

1. $|f(\boldsymbol{x})| \le C(1+\|\boldsymbol{x}\|)^{-d-N+\epsilon}$,
2. $|D^{\boldsymbol{\alpha}} f(\boldsymbol{x})| \le C(1+\|\boldsymbol{x}\|)^{-d-N-1+\epsilon}$, $|\boldsymbol{\alpha}| = 1$, and
3. $\int \boldsymbol{x}^{\boldsymbol{\beta}} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 0$, $|\boldsymbol{\beta}| < N$.

Here $C > 0$ and $0 \le \epsilon < 1$ are arbitrary constants, and the class index $N \ge 1$ is an integer. The first two properties address decay and smoothness, while the third condition guarantees vanishing moments.

3. Riesz transform of wavelets in $L_N(\mathbf{R}^d)$. This section comprises the heart of the paper, as we establish results showing how the component Riesz transforms preserve decay and vanishing moments for functions of the above class. In particular, we will show that the transform $\mathcal{R}_i f$ of a function $f \in L_N(\mathbf{R}^d)$ has a rate of decay comparable to f and has the same number of vanishing moments. Significantly, these theorems are formulated in a way that allows them to be applied iteratively so that we may derive results for the higher-order transforms in the next section. However, as the transform is given in terms of a singular integral, before estimating its decay, we need to show that it is well defined.

Proposition 3.1. *For all $f \in L_N(\mathbf{R}^d)$, $\mathcal{R}_i f$ is well defined, continuous, and bounded by*

$$C\left(\sup_{\boldsymbol{t}\in\mathbf{R}^d} \|\nabla f(\boldsymbol{t})\| + \|f\|_1\right)$$

for some $C > 0$.

Proof. We can see that the domain of the singular integral defining $\mathcal{R}_i f$ is unbounded, and we have an additional difficulty in a neighborhood of the point $\boldsymbol{x}$. We shall apply different techniques to handle these problems, so we split the integral into two terms. For a fixed $\boldsymbol{x}$, we have

$$\lim_{\epsilon\to 0}\int_{\|\boldsymbol{x}-\boldsymbol{y}\|>\epsilon} f(\boldsymbol{y})\frac{x_i-y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}}\mathrm{d}\boldsymbol{y} = \lim_{\epsilon\to 0}\int_{1>\|\boldsymbol{x}-\boldsymbol{y}\|>\epsilon} f(\boldsymbol{y})\frac{x_i-y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}}\mathrm{d}\boldsymbol{y} + \int_{\|\boldsymbol{x}-\boldsymbol{y}\|>1} f(\boldsymbol{y})\frac{x_i-y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}}\mathrm{d}\boldsymbol{y}.$$

Using the fact that the Riesz kernel is odd and the inequality

$$\frac{|f(\boldsymbol{y})-f(\boldsymbol{x})|}{\|\boldsymbol{x}-\boldsymbol{y}\|}\frac{|x_i-y_i|}{\|\boldsymbol{x}-\boldsymbol{y}\|}\frac{1}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d-1}} \le \sup_{\boldsymbol{t}\in\mathbf{R}^d}\|\nabla f(\boldsymbol{t})\|\frac{1}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d-1}} \tag{3.1}$$

holds, we can write

$$\int_{1>\|\boldsymbol{x}-\boldsymbol{y}\|>\epsilon} f(\boldsymbol{y})\frac{x_i-y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}}\mathrm{d}\boldsymbol{y} = \int_{1>\|\boldsymbol{x}-\boldsymbol{y}\|>\epsilon} (f(\boldsymbol{y})-f(\boldsymbol{x}))\frac{x_i-y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}}\mathrm{d}\boldsymbol{y} \tag{3.2}$$

and apply the dominated convergence theorem, which shows that the first term is convergent. This also shows that the first term is bounded by $C \sup_{\boldsymbol{t}\in\mathbf{R}^d} \|\nabla f(\boldsymbol{t})\|$. For the second term, we have

$$\frac{|x_i - y_i|}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}} \leq 1,$$

which implies that it is bounded by $\|f\|_1$.

To finish the proof, we must show that $\mathcal{R}_i f$ is continuous, so we fix $\boldsymbol{x} \in \mathbf{R}^d$ and $\gamma > 0$. Continuity can then be proved by showing that there are finite constants $C > 0$ and $\delta > 0$ such that $|\mathcal{R}_i f(\boldsymbol{x}) - \mathcal{R}_i f(\boldsymbol{z})| < C\gamma$ for $\|\boldsymbol{x}-\boldsymbol{z}\| < \delta$. To that end, consider the transform of f written as

$$\begin{aligned}\mathcal{R}_i f(\boldsymbol{x}) = \lim_{\epsilon\to 0} &\int_{\epsilon' > \|\boldsymbol{x}-\boldsymbol{y}\| > \epsilon} f(\boldsymbol{y}) \frac{x_i - y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}} \mathrm{d}\boldsymbol{y} \\ &+ \int_{\|\boldsymbol{x}-\boldsymbol{y}\| > \epsilon'} f(\boldsymbol{y}) \frac{x_i - y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}} \mathrm{d}\boldsymbol{y}\end{aligned}$$

for $\epsilon' > \epsilon > 0$. Using the same techniques as in (3.1) and (3.2), we can see that ϵ' should be chosen so that the absolute value of the first integral is smaller than γ. Therefore,

$$|\mathcal{R}_i f(\boldsymbol{x}) - \mathcal{R}_i f(\boldsymbol{z})| \leq 2\gamma + \int_{\|\boldsymbol{x}-\boldsymbol{y}\| > \epsilon'} |f(\boldsymbol{y})| \left| \frac{x_i - y_i}{\|\boldsymbol{x}-\boldsymbol{y}\|^{d+1}} - \frac{z_i - y_i}{\|\boldsymbol{z}-\boldsymbol{y}\|^{d+1}} \right| \mathrm{d}\boldsymbol{y},$$

and since the Riesz kernel is uniformly continuous[1] away from its singularity, there is some $\delta > 0$ (depending on ϵ') for which

$$\begin{aligned}|\mathcal{R}_i f(\boldsymbol{x}) - \mathcal{R}_i f(\boldsymbol{z})| &\leq \gamma \left(2 + \int_{\|\boldsymbol{x}-\boldsymbol{y}\| > \epsilon'} |f(\boldsymbol{y})| \, \mathrm{d}\boldsymbol{y} \right) \\ &\leq \gamma \left(2 + \|f\|_1 \right)\end{aligned}$$

when $\|\boldsymbol{x}-\boldsymbol{z}\| < \delta$. ∎

Now that we have established the fact that the transform is well defined, our focus is shifted to the main result concerning decay. Our approach is similar to that used by Adams in [2] in order to bound weighted L^p-norms of Riesz transforms, and as in that paper, a key property is the number of vanishing moments. As there are different issues with the singular integral depending on the region of $\mathbf{R}^d$, we divide it into three pieces and bound each separately. One of these regions will be handled by using the vanishing moment condition on f and approximating the shifted version of the Riesz kernel $R(\boldsymbol{x} - \cdot)$ by its Taylor expansion at $\mathbf{0}$. Therefore, for a fixed $\boldsymbol{x}$, we denote the degree $N-1$ Taylor polynomial approximant to $R(\boldsymbol{x} - \cdot)$ by $P_N(\boldsymbol{x}, \cdot)$ and define $\Phi_N(\boldsymbol{x}, \boldsymbol{y}) := R(\boldsymbol{x}-\boldsymbol{y}) - P_N(\boldsymbol{x}, \boldsymbol{y})$. With this notation, we have

$$P_N(\boldsymbol{x}, \boldsymbol{y}) + \Phi_N(\boldsymbol{x}, \boldsymbol{y}) = \sum_{|\boldsymbol{\alpha}| \leq N-1} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} R(\boldsymbol{x}) \boldsymbol{y}^{\boldsymbol{\alpha}} + \sum_{|\boldsymbol{\alpha}| = N} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}} R(\boldsymbol{x} - \boldsymbol{\xi}) \boldsymbol{y}^{\boldsymbol{\alpha}},$$

[1] This is because it is differentiable away from the singularity, and its gradient is bounded.

where $\boldsymbol{\xi}$ is a point on the line segment from $\mathbf{0}$ to $\boldsymbol{y}$. Two estimates that will be used in the proof are

$$\left|D^{\boldsymbol{\beta}} R(\boldsymbol{x})\right| \leq C\|\boldsymbol{x}\|^{-d-|\boldsymbol{\beta}|} \tag{3.3}$$

for $\boldsymbol{x} \neq \mathbf{0}$ and

$$|R(\boldsymbol{x}-\boldsymbol{y}) - P_N(\boldsymbol{x},\boldsymbol{y})| \leq C\|\boldsymbol{x}\|^{-(d+N)}\|\boldsymbol{y}\|^N \tag{3.4}$$

for $\|\boldsymbol{y}\| \leq \|\boldsymbol{x}\|/2$. The estimate (3.3) can be verified by differentiating the Riesz kernel $R(\boldsymbol{x}) = x_i/\|\boldsymbol{x}\|^{d+1}$, and (3.4) follows by applying (3.3) to the remainder in Taylor's theorem.

Theorem 3.2. *If $f \in L_N(\mathbf{R}^d)$, then the component transforms of f*

$$\mathcal{R}_i f(\boldsymbol{x}) = C_d \lim_{\epsilon' \to 0} \int_{\|\boldsymbol{x}-\boldsymbol{y}\|>\epsilon'} f(\boldsymbol{y})\Phi_N(\boldsymbol{x}-\boldsymbol{y})\mathrm{d}\boldsymbol{y}$$

satisfy the bound

$$|\mathcal{R}_i f(\boldsymbol{x})| \leq C(1+\|\boldsymbol{x}\|)^{-d-N+\epsilon+\delta}$$

for some $0 < \delta < 1-\epsilon$.

Proof. The upper bound was proved in Proposition 3.1, so it suffices to verify that $\mathcal{R}_i f(\boldsymbol{x})$ has sufficient decay for $\|\boldsymbol{x}\|$ large. In order to prove the bound, we divide the domain of the integral defining $\mathcal{R}_i f$ into the regions

$$\begin{aligned} D_1 &= \{\|\boldsymbol{y}\| < \|\boldsymbol{x}\|/2\}, \\ D_2 &= \{\|\boldsymbol{y}-\boldsymbol{x}\| \geq \|\boldsymbol{x}\|/2, \|\boldsymbol{y}\| \geq \|\boldsymbol{x}\|/2\}, \\ D_3 &= \{\|\boldsymbol{y}-\boldsymbol{x}\| < \|\boldsymbol{x}\|/2\}. \end{aligned}$$

The regions D_1 and D_3 are disjoint open balls centered at the origin and $\boldsymbol{x}$, respectively, and D_2 is the complement of their union. Since we assume $\|\boldsymbol{x}\| > 0$, these sets form a partition of $\mathbf{R}^d$. First, we consider the region D_1. Using (3.4) we get

$$\int_{D_1} |f(\boldsymbol{y})\Phi_N(\boldsymbol{x},\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} \leq C\int_{D_1} |f(\boldsymbol{y})|\,\|\boldsymbol{y}\|^N\|\boldsymbol{x}\|^{-d-N}\mathrm{d}\boldsymbol{y},$$

and since $2 < \|\boldsymbol{x}\|/\|\boldsymbol{y}\|$, the inequality

$$\int_{D_1} |f(\boldsymbol{y})\Phi_N(\boldsymbol{x},\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} \leq C\|\boldsymbol{x}\|^{-d-N+\epsilon+\delta}\left\|\|\cdot\|^{N-\epsilon-\delta}f(\cdot)\right\|_1$$

holds for any positive δ. Furthermore, due to the decay of f, the norm in the right-hand side of the above equation is finite.

Next, consider the region D_2, where we may apply (3.3) to obtain

$$\begin{aligned} \int_{D_2} |f(\boldsymbol{y})\Phi_N(\boldsymbol{x},\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} &\leq C\int_{D_2} |f(\boldsymbol{y})|\left(\|\boldsymbol{x}\|^{-d} + \sum_{k=0}^{N-1}\|\boldsymbol{y}\|^k\|\boldsymbol{x}\|^{-d-k}\right)\mathrm{d}\boldsymbol{y} \\ &\leq C\|\boldsymbol{x}\|^{-d}\int_{D_2} |f(\boldsymbol{y})|\left(\sum_{k=0}^{N-1}\frac{\|\boldsymbol{y}\|^k}{\|\boldsymbol{x}\|^k}\right)\mathrm{d}\boldsymbol{y}, \end{aligned}$$

and note that for $k \leq N-1$, we have $N-(\epsilon+\delta)-k > 0$, since $0 \leq \epsilon+\delta < 1$. Additionally, $2\|\boldsymbol{y}\| \geq \|\boldsymbol{x}\|$ in this region, so we multiply each term of the sum by $(2\|\boldsymbol{y}\|/\|\boldsymbol{x}\|)^{N-(\epsilon+\delta)-k}$ to obtain the required bound:

$$\begin{aligned}\int_{D_2} |f(\boldsymbol{y})\Phi_N(\boldsymbol{x},\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} &\leq C\|\boldsymbol{x}\|^{-d}\int_{D_2}|f(\boldsymbol{y})|\left(\frac{\|\boldsymbol{y}\|}{\|\boldsymbol{x}\|}\right)^{N-\epsilon-\delta}\mathrm{d}\boldsymbol{y}\\ &\leq C\|\boldsymbol{x}\|^{-d-N+\epsilon+\delta}\left\|\|\cdot\|^{N-\epsilon-\delta}f(\cdot)\right\|_1.\end{aligned}$$

Finally, the region D_3 contains the singularity of the Riesz kernel, and we further divide the integral,

$$\begin{aligned}\lim_{\epsilon'\to 0}\int_{D_3\setminus B(\boldsymbol{x},\epsilon')} f(\boldsymbol{y})\Phi_N(\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{y} &= \lim_{\epsilon'\to 0}\int_{D_3\setminus B(\boldsymbol{x},\epsilon')} f(\boldsymbol{y})R(\boldsymbol{x}-\boldsymbol{y})\mathrm{d}\boldsymbol{y}\\ &\quad -\int_{D_3} f(\boldsymbol{y})P_N(\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{y},\end{aligned}$$

where $B(\boldsymbol{x},\epsilon')$ denotes the ball of radius ϵ' centered at $\boldsymbol{x}$. The second part is bounded using techniques similar to those used in the region D_2:

$$\begin{aligned}\int_{D_3}|f(\boldsymbol{y})P_N(\boldsymbol{x},\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} &\leq C\int_{D_3}|f(\boldsymbol{y})|\sum_{k=0}^{N-1}\|\boldsymbol{y}\|^k\|\boldsymbol{x}\|^{-d-k}\mathrm{d}\boldsymbol{y}\\ &\leq C\int_{D_3}|f(\boldsymbol{y})|\,\|\boldsymbol{x}\|^{-d}\left(\frac{\|\boldsymbol{y}\|}{\|\boldsymbol{x}\|}\right)^{N-1}\mathrm{d}\boldsymbol{y}\\ &\leq C\|\boldsymbol{x}\|^{-N-d+\epsilon+\delta}\left\|\|\cdot\|^{N-\epsilon-\delta}f(\cdot)\right\|_1.\end{aligned}$$

The first part is bounded using the mean value theorem along with the estimate for the decay of the derivatives of f:

$$\begin{aligned}\left|\int_{D_3\setminus B(\boldsymbol{x},\epsilon')} f(\boldsymbol{y})R(\boldsymbol{x}-\boldsymbol{y})\mathrm{d}\boldsymbol{y}\right| &= \left|\int_{D_3\setminus B(\boldsymbol{x},\epsilon')}\frac{(f(\boldsymbol{x})-f(\boldsymbol{y}))}{\|\boldsymbol{x}-\boldsymbol{y}\|}\frac{(x_i-y_i)}{\|\boldsymbol{x}-\boldsymbol{y}\|^d}\mathrm{d}\boldsymbol{y}\right|\\ &\leq C\sup_{\boldsymbol{t}\in B(\boldsymbol{x},\|\boldsymbol{x}\|/2)}\|\nabla f(\boldsymbol{t})\|\int_{D_3}\|\boldsymbol{x}-\boldsymbol{y}\|^{-d+1}\mathrm{d}\boldsymbol{y}\\ &\leq C(1+\|\boldsymbol{x}\|)^{-d-N+\epsilon}.\end{aligned}$$

As this bound is independent of ϵ', the result follows. ■

As previously stated, the Riesz transform preserves both decay and vanishing moments, and both are vital for deriving decay estimates for the higher-order Riesz transforms. The former property was the subject of the last theorem, and we now address the latter; i.e., we will show that for any $f \in L_N(\mathbf{R}^d)$, $\mathcal{R}_i f$ has vanishing moments up to degree $N-1$. Basically, the idea is to consider both the Fourier and spatial domain formulations concurrently. To wit, vanishing moments can be interpreted both as the order of the zero of a function at the origin in the Fourier domain and as an integral.

Proposition 3.3. *For $f \in L_N(\mathbf{R}^d)$ and $|\boldsymbol{\beta}| < N$, we have*

$$D^{\boldsymbol{\beta}} \widehat{f}(\boldsymbol{\omega}) = o(\|\boldsymbol{\omega}\|^{N-1-|\boldsymbol{\beta}|}) \quad \text{as } \boldsymbol{\omega} \to \mathbf{0}.$$

Proof. The decay estimate on f implies that $(\cdot)^{\boldsymbol{\beta}} f$ is in $L^1(\mathbf{R}^d)$ for $|\beta| < N$, so $\widehat{f}$ has continuous derivatives up to order $N-1$ and

$$D^{\boldsymbol{\beta}} \widehat{f} = C[(\cdot)^{\boldsymbol{\beta}} f]^{\wedge};$$

cf. [20, Theorem 1.1.7]. The right-hand side can be expressed as the integral

$$D^{\boldsymbol{\beta}} \widehat{f}(\boldsymbol{\omega}) = C \int \boldsymbol{x}^{\boldsymbol{\beta}} f(\boldsymbol{x}) e^{j\boldsymbol{x}\cdot\boldsymbol{\omega}} \mathrm{d}\boldsymbol{x},$$

and the vanishing moment condition on f implies that $D^{\boldsymbol{\beta}} \widehat{f}(\mathbf{0}) = 0$ for $|\boldsymbol{\beta}| < N$. Therefore, Taylor's theorem implies that

$$D^{\boldsymbol{\beta}} \widehat{f}(\boldsymbol{\omega}) = \sum_{|\boldsymbol{\alpha}|=N-1-|\boldsymbol{\beta}|} \frac{1}{\boldsymbol{\alpha}!} D^{\boldsymbol{\alpha}+\boldsymbol{\beta}} \widehat{f}(\boldsymbol{\xi}) \boldsymbol{\omega}^{\boldsymbol{\alpha}}$$

for some $\boldsymbol{\xi}$ on the line segment from $\mathbf{0}$ to $\boldsymbol{\omega}$. The result now follows from the fact that $D^{\boldsymbol{\alpha}+\boldsymbol{\beta}} \widehat{f}(\boldsymbol{\xi}) \to 0$ as $\boldsymbol{\xi} \to \mathbf{0}$. ∎

Theorem 3.4. *Given any $f \in L_N(\mathbf{R}^d)$, $\mathcal{R}_i f$ has vanishing moments up to degree $N-1$.*

Proof. The proof relies on considering two representations of $\widehat{\mathcal{R}_i f}$. First we have

$$\widehat{\mathcal{R}_i f}(\boldsymbol{\omega}) = C \int \mathcal{R}_i f(\boldsymbol{x}) e^{j\boldsymbol{x}\cdot\boldsymbol{\omega}} \mathrm{d}\boldsymbol{x},$$

so

$$D^{\boldsymbol{\beta}} \widehat{\mathcal{R}_i f}(\boldsymbol{\omega}) = C \int \boldsymbol{x}^{\boldsymbol{\beta}} \mathcal{R}_i f(\boldsymbol{x}) e^{j\boldsymbol{x}\cdot\boldsymbol{\omega}} \mathrm{d}\boldsymbol{x}.$$

We also know that

$$\widehat{\mathcal{R}_i f}(\boldsymbol{\omega}) = \frac{\omega_i}{\|\boldsymbol{\omega}\|} \widehat{f}(\boldsymbol{\omega}),$$

and hence

$$D^{\boldsymbol{\beta}} \widehat{\mathcal{R}_i f}(\mathbf{0}) = 0$$

for $|\boldsymbol{\beta}| < N$ by Proposition 3.3. Putting the two together, we have

$$0 = D^{\boldsymbol{\beta}} \widehat{\mathcal{R}_i f}(\mathbf{0}) = C \int \boldsymbol{x}^{\boldsymbol{\beta}} \mathcal{R}_i f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \qquad ∎$$

Let us conclude this section by considering some examples. To see how our results can be applied, we will first examine the Hermitian wavelets in one dimension, which we recall are derivatives of the Gaussian,

$$\psi_n(t) = C_n \frac{d^n}{dt^n} e^{-t^2/2},$$

where C_n is a normalization constant. Clearly, these wavelets are infinitely smooth and decay faster than any polynomial. Therefore, we must look at their vanishing moments to determine

the spaces L_N in which they lie. Being derivatives of the Gaussian, integration by parts shows that ψ_n has n vanishing moments, so $\psi_n \in L_n(\mathbf{R})$. Applying the results of this section, we conclude that the Riesz transform of ψ_n, denoted by $\mathcal{R}\psi_n$, satisfies $|\mathcal{R}\psi_n(x)| \leq C(1+|x|)^{-n}$ and $\mathcal{R}\psi_n$ has n vanishing moments. Note that since we are interested only in the essential decay of the Riesz transform, we have bounded the term $|x|^{\delta+\epsilon}$ by $|x|$ in our estimate. However, the more precise decay rate stated in Theorem 3.2 will be important for showing how decay is preserved for higher-order Riesz transforms.

As a second example, consider the two-dimensional Poisson kernel

$$p(\boldsymbol{x}; s) = \frac{s}{2\pi(s^2 + \|\boldsymbol{x}\|^2)^{3/2}},$$

which was studied by Felsberg and Sommer [6, 8]. The components of the Riesz transform of this function are computed to be

$$\mathcal{R}_i p(\boldsymbol{x}; s) = C_d \frac{x_i}{(s^2 + \|\boldsymbol{x}\|^2)^{3/2}},$$

which have decay comparable to $\|\boldsymbol{x}\|^{-2}$ in the direction x_i. Here, we can again see that while the Poisson kernel is smooth, its Riesz transform has slower decay due to the fact that $p(\boldsymbol{x}; s)$ has no vanishing moments.

4. Higher-order Riesz transforms. In the previous section, we showed that a component transform of a function in $L_N(\mathbf{R}^d)$ has a rate of decay similar to that of the original function and a comparable number of vanishing moments. With this in mind, we would like to extend these results to higher-order Riesz transforms by applying the results for the single order transform in an iterated fashion. In order to accomplish this, we will need to impose conditions which are stronger than those required for $L_N(\mathbf{R}^d)$. To make this precise, we recall the definition of the Sobolev spaces on $\mathbf{R}^d$.

Definition 4.1. *Let $H^s(\mathbf{R}^d)$ denote the L^2 Sobolev space of order $s \geq 0$ on $\mathbf{R}^d$, i.e.,*

$$H^s(\mathbf{R}^d) = \left\{ f \in L^2(\mathbf{R}^d) : \widehat{f}(\cdot)(1 + \|\cdot\|_2^2)^{s/2} \in L^2(\mathbf{R}^d) \right\}.$$

The additional properties that we required of the original wavelet are given in terms of its derivatives. In particular, the derivatives of the wavelet should also have decay and vanishing moments.

Definition 4.2. *For $N \in \mathbf{N}$ and $M \in \mathbf{N}$ we define the space $L_{N,M}(\mathbf{R}^d)$ to be the collection of $f : \mathbf{R}^d \to \mathbf{R}$ satisfying the following:*

1. *$f \in H^s(\mathbf{R}^d) \cap C(\mathbf{R}^d)$ for some $s > M + d/2$.*
2. *There are constants $C > 0$ and $0 \leq \epsilon < 1$ such that for any $0 \leq |\boldsymbol{\alpha}| \leq M$ we have*
   (a) *$|D^{\boldsymbol{\alpha}} f(\boldsymbol{x})| \leq C(1 + \|\boldsymbol{x}\|)^{-d-N-|\boldsymbol{\alpha}|+\epsilon}$ for $0 \leq |\boldsymbol{\alpha}| \leq M$, and*
   (b) *$\int \boldsymbol{x}^{\boldsymbol{\beta}} D^{\boldsymbol{\alpha}} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 0$ for $0 \leq |\boldsymbol{\alpha}| \leq M - 1$ and $|\boldsymbol{\beta}| < N + |\boldsymbol{\alpha}|$.*

Notice that the first property in the definition implies $L_{N,M}(\mathbf{R}^d)$ is contained in $C^M(\mathbf{R}^d)$ by the Sobolev embedding theorem. Our definition also guarantees that $L_{N,M}(\mathbf{R}^d)$ is a subset of $L_N(\mathbf{R}^d)$ for $M \geq 1$, and we will show that the component transforms map $L_{N,M}(\mathbf{R}^d)$ to $L_{N,M-1}(\mathbf{R}^d)$. Hence the iterated transforms of a function in $L_{N,M}(\mathbf{R}^d)$ will have a rate of

decay comparable to that of the single transform of a function in $L_N(\mathbf{R}^d)$. The proof relies on the iterated application of Theorems 3.2 and 3.4; however, we must first establish that the transforms commute with differentiation.

Lemma 4.3. *If $f \in L_{N,M}(\mathbf{R}^d)$ for some $M > 1$, then $\mathcal{R}_i D^{\boldsymbol{\alpha}} f = D^{\boldsymbol{\alpha}} \mathcal{R}_i f$ for any $0 < |\boldsymbol{\alpha}| \leq M - 1$.*

Proof. First, since $f \in H^s$ for some $s > M + d/2$, the Sobolev embedding theorem implies $f \in C^M(\mathbf{R}^d)$. Additionally,

$$\widehat{\mathcal{R}_i f} = \frac{\omega_i}{\|\boldsymbol{\omega}\|} \widehat{f},$$

so $\mathcal{R}_i f$ is also in H^s and consequently in $C^M(\mathbf{R}^d)$. Therefore, for any fixed $0 < |\boldsymbol{\alpha}| \leq M - 1$, $\mathcal{R}_i D^{\boldsymbol{\alpha}} f$ and $D^{\boldsymbol{\alpha}} \mathcal{R}_i f$ are continuous functions. Furthermore, they are equivalent as distributions since they have the same Fourier transform $\omega_i / \|\boldsymbol{\omega}\| \boldsymbol{\omega}^{\boldsymbol{\alpha}} \widehat{f}$. Therefore, they are equivalent as functions [11, Theorem, p.27]. ∎

Combining this lemma with the results of the previous section, we can see how the defining properties of $L_{N,M}$ ensure decay of the higher-order Riesz transforms, as we explicitly show in the next few results.

Lemma 4.4. *If $f \in L_{N,M}(\mathbf{R}^d)$ for some $M > 1$, then $\mathcal{R}_i f \in L_{N,M-1}(\mathbf{R}^d)$.*

Proof. By Lemma 4.3, it suffices to verify the second property in the definition of $L_{N,M-1}(\mathbf{R}^d)$. To begin, fix a multi-index $\boldsymbol{\alpha}$ with $0 \leq |\boldsymbol{\alpha}| \leq M - 1$. Since $f \in L_{N,M}(\mathbf{R}^d)$, we know that

1. $|D^{\boldsymbol{\alpha}} f(\boldsymbol{x})| \leq C(1 + \|\boldsymbol{x}\|)^{-d-N-|\boldsymbol{\alpha}|+\epsilon}$,
2. $|D^{\boldsymbol{\alpha}+\boldsymbol{\gamma}} f(\boldsymbol{x})| \leq C(1 + \|\boldsymbol{x}\|)^{-d-N-|\boldsymbol{\alpha}|-1+\epsilon}$ for $|\boldsymbol{\gamma}| = 1$, and
3. $D^{\boldsymbol{\alpha}} f$ has vanishing moments up to degree $N - 1 + |\boldsymbol{\alpha}|$.

Theorem 3.2 then implies that

$$D^{\boldsymbol{\alpha}} \mathcal{R}_i f(\boldsymbol{x}) = \mathcal{R}_i D^{\boldsymbol{\alpha}} f(\boldsymbol{x}) \leq C(1 + \|\boldsymbol{x}\|)^{-d-N-|\boldsymbol{\alpha}|+\epsilon+\delta}$$

for some $0 < \delta < 1 - \epsilon$. Additionally, Theorem 3.4 implies that $D^{\boldsymbol{\alpha}} \mathcal{R}_i f(\boldsymbol{x})$ has vanishing moments up to degree $N - 1 + |\boldsymbol{\alpha}|$. ∎

Theorem 4.5. *If $f \in L_{N,M}(\mathbf{R}^d)$, then for some $0 < \epsilon < 1$ and any $1 \leq |\boldsymbol{\alpha}| \leq M$ we have $\mathcal{R}^{\boldsymbol{\alpha}} f(\boldsymbol{x}) \leq C(1 + \|\boldsymbol{x}\|)^{-d-N+\epsilon}$, and $\mathcal{R}^{\boldsymbol{\alpha}} f(\boldsymbol{x})$ has vanishing moments up to degree $N - 1$.*

To illustrate this result, we consider a two-dimensional analogue of the Hermitian wavelets,

$$\psi_n(\boldsymbol{x}) = C_n D^{\boldsymbol{\beta}} e^{-\|\boldsymbol{x}\|^2/2},$$

where $\beta_1 = \beta_2 = 2n$ for some positive integer n. As before, the number of vanishing moments is the limiting factor when classifying ψ_n in terms of the spaces $L_{N,M}$. Considering their Fourier transforms, we can see that ψ_n is in $L_{n,n}$. Therefore, Theorem 4.5 implies that the nth-order Riesz transform of ψ_n, denoted by $\mathcal{R}^n \psi_n$, decays at least as quickly as $(1 + \|\boldsymbol{x}\|)^{-n}$. Moreover, notice that if ψ_n had infinitely many vanishing moments, this theorem would guarantee a rate of decay faster than any polynomial. In fact, the Riesz transform preserves the collection of Schwartz class functions with infinitely many vanishing moments, as stated formally in the following corollary.

Corollary 4.6. *If ψ is a Schwartz class function and all of its moments vanish, then $\mathcal{R}_i \psi$ satisfies the same properties.*

5. Application to Simoncelli wavelets. As we have seen, spatial decay of the Riesz transform is guaranteed by assuming smoothness, vanishing moments, and decay of the original function. There are several examples of wavelets that are bandlimited and whose Fourier transform is zero in a neighborhood of the origin (cf. [22]). The one property that many wavelet constructions lack is fast spatial decay. This is a result of constructing wavelets in the Fourier domain using a nonsmooth cut-off function. One way to see this is to consider a Fourier integral representation of a wavelet ψ:

$$\psi(x) = \frac{1}{(2\pi)^{1/2}} \int_{\mathbf{R}} \widehat{\psi}(\omega) e^{j\omega x} \mathrm{d}\omega.$$

If $\widehat{\psi}$ has n continuous derivatives and compact support, then we can integrate by parts to get the decay estimate

$$|\psi(x)| = \frac{1}{(2\pi)^{1/2}} \frac{\left\| \widehat{\psi}^{(n)} \right\|_1}{|x|^n}.$$

However, a discontinuity of a lower-order derivative of $\widehat{\psi}$ can produce terms with slower decay. This estimate also shows how the size of the derivatives of $\widehat{\psi}$ play a role in the decay of ψ. To give a specific example and show how decay can be improved, we consider the Simoncelli wavelets [18], defined radially in the Fourier domain by

$$\widehat{\psi}(\omega) = \begin{cases} \cos(\frac{\pi}{2} \log_2(\frac{2\omega}{\pi})), & \omega \in (\frac{\pi}{4}, \pi], \\ 0 & \text{otherwise.} \end{cases}$$

Note that this wavelet has vanishing moments of all degrees and finite Sobolev norm of all orders. However, the lack of smoothness in the Fourier domain causes ψ to have slow decay. With Theorem 4.5 as a guideline, we propose a modification to improve the decay of its higher Riesz transforms.

If we write $\widehat{\psi}$ as $(\chi_{[-1,1]} \cos(\frac{\pi}{2}\cdot)) \circ \log_2(\frac{2}{\pi}\cdot)$, where $\circ$ denotes function composition, it is clear that we can create a smooth analogue if we replace the indicator function χ by a smooth Meyer window. In particular, let θ_ϵ be a function satisfying the following:

1. θ_ϵ is identically equal to 1 on the interval $[-1+\epsilon, 1-\epsilon]$.
2. θ_ϵ is equal to 0 outside $[-1-\epsilon, 1+\epsilon]$.
3. θ_ϵ has at least r continuous derivatives for some $r \in \mathbb{N}$.
4. $\theta_\epsilon(x)^2 + \theta_\epsilon(x-2)^2 = 1$ for $0 \le x \le 2$.

Now, if we merely replace the indicator function with θ_ϵ, we will have increased the support of our wavelet in the Fourier domain, and the basis it generates might no longer be a tight frame. To compensate, we rescale so that support is maintained within $[-\pi, \pi]^d$. Let us denote the modified wavelet by ψ'_ϵ, which in the Fourier domain is defined radially by

$$\left(\theta_\epsilon \cos\left(\frac{\pi}{2}\cdot\right)\right) \circ \log_2\left(\frac{2^{1+\epsilon}}{\pi}\cdot\right).$$

The decay of this new function and its transforms can be determined by the results of the previous sections.

Table 1

Examples of finite smoothness Meyer windows: Replacing G by G_n in (5.1) and using the resulting $H_{n,\epsilon}$ in (5.2), we obtain a $\theta_{n,\epsilon}$ in $C^n(\mathbf{R})$.

n	$G_n(x)$
3	$\frac{35\pi}{64}(-\frac{1}{7}x^7+\frac{3}{5}x^5-x^3+x+\frac{16}{35})$
4	$\frac{315\pi}{512}(\frac{1}{9}x^9-\frac{4}{7}x^7+\frac{6}{5}x^5-\frac{4}{3}x^3+x+\frac{128}{315})$
5	$\frac{693\pi}{1024}(-\frac{1}{11}x^{11}+\frac{5}{9}x^9-\frac{10}{7}x^7+2x^5-\frac{5}{3}x^3+x+\frac{256}{693})$

Proposition 5.1. *The value of r defining the smoothness of θ_ϵ will determine the decay of ψ'_ϵ, and the higher Riesz transforms of ψ'_ϵ will maintain a similar rate of decay.*

Here, we should point out that one of the nice properties of the Simoncelli wavelets is that they generate a tight frame, and importantly, our modification does not nullify this property.

Proposition 5.2. *The wavelet basis constructed from the modified Simoncelli wavelet ψ'_ϵ is a tight frame.*

Proof. See the appendix. ∎

We conclude now with a discussion about the construction of the Meyer windows θ_ϵ. The starting point is a smooth, nonnegative bump function g that is supported in $[-1,1]$ and whose definite integral is $\pi/2$. From this function, we define

$$G(\omega) := \int_{-\infty}^{\omega} g(t)\mathrm{d}t$$

and

$$H_\epsilon(\omega) := G\left(\frac{\omega+1}{\epsilon}\right) - \frac{\pi}{2} + G\left(\frac{\omega-1}{\epsilon}\right). \tag{5.1}$$

The Meyer window can then be constructed as

$$\theta_\epsilon(\omega) = \cos(H_\epsilon(\omega)). \tag{5.2}$$

Indeed, one can verify that for $0 \le \omega \le 2$ we have

$$\begin{aligned}\theta_\epsilon(\omega)^2 + \theta_\epsilon(\omega-2)^2 &= \cos^2(H_\epsilon(\omega)) + \cos^2(H_\epsilon(\omega-2)) \\ &= \cos^2(H_\epsilon(\omega)) + \cos^2(H_\epsilon(\omega)-\pi/2) \\ &= \cos^2(H_\epsilon(\omega)) + \sin^2(H_\epsilon(\omega)) \\ &= 1.\end{aligned}$$

In the case of finite smoothness, examples can be found by setting $g_n(x) = C(1-x^2)^n$ on $[-1,1]$ and 0 otherwise. For some particular examples, see Table 1. In fact, the functions in the table are also derived in [10]; however, the authors' construction is different from ours. Notice that each such G_n has n continuous derivatives, so the corresponding $\theta_{n,\epsilon}$ and $\widehat{\psi}'_{n,\epsilon}$ will also have n continuous derivatives, which translates into decay of the wavelet in the spatial domain.

In order to construct an infinitely smooth θ, we take another approach, as explained in [10]. Instead of starting with a bump function, we directly define the function G used in (5.1). Specifically, a function G with the required properties is given by

$$G(\omega) = \frac{\pi}{2} \frac{\lambda(\omega + 1)}{\lambda(\omega + 1) + \lambda(1 - \omega)},$$

where

$$\lambda(\omega) = \begin{cases} e^{-\omega^{-2}}, & \omega > 0, \\ 0 & \text{otherwise}, \end{cases}$$

and the result of this construction is a Meyer window with infinite smoothness.

6. Conclusion. In this paper, we have derived decay estimates for higher-order Riesz transforms of wavelets, where the conditions are given in terms of smoothness, decay, and vanishing moments. In the case of a single Riesz transform, it was shown that the decay rate of the original function is essentially maintained by the component transforms, and this result is likely to be optimal. These results instill confidence in recently proposed steerable wavelet constructions, as the Riesz transform plays a large role in the design. Additionally, they point out the importance of specifying primal wavelets with a large number of vanishing moments, a fast rate of decay, and a sufficient amount of smoothness. In particular, our results indicate that having many vanishing moments is not sufficient to guarantee decay; a matched degree of smoothness in both the space and frequency domains is also required. Optimal conditions for the decay estimates of the higher-order transforms remain an open question. By using the single-order results in an iterated fashion, we may be imposing conditions that are stronger than necessary, and taking a more direct approach may provide improved results. Given the importance of the Riesz transform in steerable wavelet constructions, we view this as a topic that warrants further investigation.

Appendix. Proof of Proposition 5.2. Due to the support of $\widehat{\psi}'_\epsilon$, it suffices to verify

$$\sum_{k \in \mathbb{Z}} \left| \widehat{\psi}'_\epsilon(2^k \omega) \right|^2 = 1$$

for $\omega > 0$. Using the definition, we have

$$\sum_{k \in \mathbb{Z}} \left| \widehat{\psi}'_\epsilon(2^k \omega) \right|^2 = \sum_{k \in \mathbb{Z}} \left| \left(\theta_\epsilon \cos\left(\frac{\pi}{2} \cdot \right) \right) \circ \log_2 \left(\frac{2^{1+\epsilon}}{\pi} 2^k \omega \right) \right|^2,$$

and setting $z = \log_2(2^{1+\epsilon}\omega/\pi)$ gives

$$\sum_{k \in \mathbb{Z}} \left| \widehat{\psi}'_\epsilon(2^k \omega) \right|^2 = \sum_{k \in \mathbb{Z}} \left| \theta_\epsilon(z + k) \cos\left(\frac{\pi}{2}(z + k) \right) \right|^2.$$

This sum is evaluated by considering two cases that depend on z. First, for $-\epsilon \le z \le \epsilon$, we have

$$\begin{aligned}\sum_{k\in\mathbb{Z}} \left|\widehat{\psi}'_\epsilon(2^k\omega)\right|^2 &= \sum_{k=-1}^{1} \left|\theta_\epsilon(z+k)\cos\left(\frac{\pi}{2}(z+k)\right)\right|^2 \\ &= \cos^2\left(\frac{\pi}{2}z\right) + \cos^2\left(\frac{\pi}{2}(z-1)\right)\theta_\epsilon(z-1)^2 \\ &\quad + \cos^2\left(\frac{\pi}{2}(z+1)\right)\theta_\epsilon(z+1)^2 \\ &= \sin^2\left(\frac{\pi}{2}(z-1)\right) \\ &\quad + \left(\theta_\epsilon(z-1)^2 + \theta_\epsilon(z+1)^2\right)\cos^2\left(\frac{\pi}{2}(z-1)\right) \\ &= 1.\end{aligned}$$

Next, for $\epsilon \le z \le 1-\epsilon$, we have

$$\begin{aligned}\sum_{k\in\mathbb{Z}} \left|\widehat{\psi}'_\epsilon(2^k\omega)\right|^2 &= \sum_{k=-1}^{0} \left|\theta_\epsilon(z+k)\cos\left(\frac{\pi}{2}(z+k)\right)\right|^2 \\ &= \cos^2\left(\frac{\pi}{2}(z-1)\right) + \cos^2\left(\frac{\pi}{2}z\right) \\ &= \sin^2\left(\frac{\pi}{2}z\right) + \cos^2\left(\frac{\pi}{2}z\right) \\ &= 1.\end{aligned}$$

REFERENCES

[1] P. Abry and P. Flandrin, *Multiresolution transient detection*, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, 1994, pp. 225–228.

[2] E. Adams, *On weighted norm inequalities for the Riesz transforms of functions with vanishing moments*, Studia Math., 78 (1984), pp. 107–153.

[3] K. N. Chaudhury and M. Unser, *On the Hilbert transform of wavelets*, IEEE Trans. Signal Process., 59 (2011), pp. 1890–1894.

[4] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[5] J. Duoandikoetxea, *Fourier Analysis*, AMS, Providence, RI, 2000.

[6] M. Felsberg, *On the design of two-dimensional polar separable filters*, in Proceedings of the 12th European Signal Processing Conference, Vienna, Austria, 2004, pp. 417–420.

[7] M. Felsberg and G. Sommer, *The monogenic signal*, IEEE Trans. Signal Process., 49 (2001), pp. 3136–3144.

[8] M. Felsberg and G. Sommer, *The monogenic scale-space: A unifying approach to phase-based image processing in scale-space*, J. Math. Imag. Vis., 21 (2004), pp. 5–26.

[9] S. L. Hahn, *Hilbert transforms in signal processing*, Artech House, Boston, 1996.

[10] S. Held, M. Storath, P. Massopust, and B. Forster, *Steerable wavelet frames based on the Riesz transform*, IEEE Trans. Image Process., 19 (2010), pp. 653–667.

[11] R. P. Kanwal, *Generalized Functions*, Academic Press, New York, 1983.

[12] H. Knutsson and M. Andersson, *Loglets: Generalized quadrature and phase for local spatio-temporal structure estimation*, in Image Analysis, Lecture Notes in Comput. Sci. 2749, Springer, Berlin, 2003.

[13] U. Köthe and M. Felsberg, *Riesz-transforms versus derivatives: On the relationship between the boundary tensor and the energy tensor*, in Scale Space and PDE Methods in Computer Vision, Lecture Notes in Comput. Sci. 3459, Springer, Berlin, 2005, pp. 179–191.

[14] K. G. Larkin, D. J. Bone, and M. A. Oldfield, *Natural demodulation of two-dimensional fringe patterns. I. General background of the spiral phase quadrature transform*, J. Opt. Soc. Amer. A, 18 (2011), pp. 1862–1870.
[15] K. G. Larkin, *Natural demodulation of two-dimensional fringe patterns. II. Stationary phase analysis of the spiral phase quadrature transform*, J. Opt. Soc. Amer. A, 18 (2011), pp. 1871–1881.
[16] G. Metikas and S. C. Olhede, *Multiple multidimensional Morse wavelets*, IEEE Trans. Signal Process., 55 (2007), pp. 921–936.
[17] T. Olson and J. DeStefano, *Wavelet localization of the Radon transform*, IEEE Trans. Signal Process., 42 (1994), pp. 2055–2067.
[18] J. Portilla and E. P. Simoncelli, *A parametric texture model based on joint statistics of complex wavelet coefficients*, Int. J. Comput. Vis., 40 (2000), pp. 49–70.
[19] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, *Shiftable multiscale transforms*, IEEE Trans. Inform. Theory, 38 (1992), pp. 587–607.
[20] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, NJ, 1971.
[21] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ, 1970.
[22] M. Unser, N. Chenouard, and D. Van De Ville, *Steerable pyramids and tight wavelet frames in $L^2(\mathbf{R}^d)$*, IEEE Trans. Image Process., 20 (2011), pp. 2705–2721.
[23] M. Unser, D. Sage, and D. Van De Ville, *Multiresolution monogenic signal analysis using the Riesz-Laplace wavelet transform*, IEEE Trans. Image Process., 18 (2009), pp. 2402–2418.
[24] M. Unser and D. Van De Ville, *Wavelet steerability and the higher-order Riesz transform*, IEEE Trans. Image Process., 19 (2010), pp. 636–652.