

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on April 28, 1998.

Tight frames of k -plane ridgelets and the problem of representing objects that are smooth away from d -dimensional singularities in \mathbf{R}^n

(X -ray transform / ridge function / wavelets)

DAVID L. DONOHO[†]

Department of Statistics, Stanford University, Stanford, CA 94305

Contributed by David L. Donoho, December 29, 1998

ABSTRACT For each pair (n, k) with $1 \leq k < n$, we construct a tight frame $(\rho_\lambda : \lambda \in \Lambda)$ for $L^2(\mathbf{R}^n)$, which we call a frame of k -plane ridgelets. The intent is to efficiently represent functions that are smooth away from singularities along k -planes in \mathbf{R}^n . We also develop tools to help decide whether k -plane ridgelets provide the desired efficient representation. We first construct a wavelet-like tight frame on the X -ray bundle $\mathcal{X}_{n,k}$ —the fiber bundle having the Grassman manifold $G_{n,k}$ of k -planes in \mathbf{R}^n for base space, and for fibers the orthocomplements of those planes. This wavelet-like tight frame is the pushout to $\mathcal{X}_{n,k}$, via the smooth local coordinates of $G_{n,k}$, of an orthonormal basis of tensor Meyer wavelets on Euclidean space $\mathbf{R}^{k(n-k)} \times \mathbf{R}^{n-k}$. We then use the X -ray isometry [Solmon, D. C. (1976) *J. Math. Anal. Appl.* 56, 61–83] to map this tight frame isometrically to a tight frame for $L^2(\mathbf{R}^n)$ —the k -plane ridgelets. This construction makes analysis of a function $f \in L^2(\mathbf{R}^n)$ by k -plane ridgelets identical to the analysis of the k -plane X -ray transform of f by an appropriate wavelet-like system for $\mathcal{X}_{n,k}$. As wavelets are typically effective at representing point singularities, it may be expected that these new systems will be effective at representing objects whose k -plane X -ray transform has a point singularity. Objects with discontinuities across hyperplanes are of this form, for $k = n - 1$.

1. Introduction

One of the most striking features of wavelet analysis is its ability to efficiently represent functions that are smooth away from point singularities. To understand this, consider the function $f_{0,\alpha}(x) = |x|^{-\alpha}w(x)$ on \mathbf{R}^n , where $w(x)$ is a smooth window of compact support and $\alpha < n/2$. Now f is smooth away from 0, and has a square-integrable singularity at the point $x = 0$. The coefficients of f in the Meyer orthonormal wavelet basis are sparse: arranging them in decreasing order of magnitude gives a sequence decaying more rapidly than any negative power of the index. In this regard, the wavelet coefficients of a point singularity behave similarly to the wavelet coefficients of a smooth function [such as $w(x)$]; the sparsity of a wavelet analysis is in a sense insensitive to the presence of point singularities.

Sparsity of the wavelet coefficients has implications for the quality of partial wavelet reconstructions. If we approximate a function by using just the m -best terms in the wavelet expansion, and if the coefficients are sparse in the sense just given, then the L^2 error of best- m -term approximation decays rapidly with m —faster than any negative power of m .

Hence, the fact that wavelet analysis of a point singularity yields sparse coefficients means that smooth functions with point singularities can be very efficiently approximated by partial wavelet reconstructions. This fact has significant implications in data compression and in statistical estimation. (Extensive references on these implications are given in refs. 1 and 2.)

In dimension $n > 1$, there is a wide range of singularity types, point singularities being just one possibility. Consider

$$f_{d,\alpha}(x) = w(x_1, \dots, x_n) \cdot (x_1^2 + \dots + x_{n-d}^2)^{-\alpha/2},$$

where w is a smooth window of compact support, $d \in \{1, \dots, n-1\}$, and $0 < \alpha < (n-d)/2$. This function has a singularity along the hyperplane $x_1 = \dots = x_{n-d} = 0$ that extends a finite distance in the d variables x_{n-d+1}, \dots, x_n . It may naturally be viewed as a singularity of dimension d . We may also naturally consider rigid motions of the argument, producing $\tilde{f}_{d,\alpha}(x) = f_{d,\alpha}(Ux+b)$, where U is a rotation of \mathbf{R}^n .

For typical functions of the type $\tilde{f}_{d,\alpha}$, $d \neq 0$, wavelets do not yield sparse coefficients as they did with $f_{0,\alpha}$. For example, in \mathbf{R}^2 , an object of type $\tilde{f}_{1,1/4}$ is easily seen to have typically at least order $O(2^j)$ standard wavelet coefficients with amplitude exceeding $2^{-3j/4}$. So the m th largest wavelet coefficient of such an object is often $\geq c \cdot m^{-3/4}$ for $c > 0$; this is a much poorer decay than the one we saw earlier in the case of point singularities, where the decay was faster than any negative power of m . In consequence, m -term wavelet reconstructions do not approximate such objects with the kind of efficiency we saw earlier in the case of point singularities. We can formulate this conclusion more boldly by saying that wavelets do not efficiently approximate edges in \mathbf{R}^2 . Similar statements hold in higher dimensions; wavelets do not efficiently approximate discontinuities across surfaces in \mathbf{R}^3 or singularities along curves in \mathbf{R}^3 .

We summarize this as follows: wavelets efficiently represent 0-dimensional singularities, but not d -dimensional singularities, for $d = 1, \dots, n-1$ in dimension n . This leads very naturally to the following problem:

PROBLEM (n, d): Representation of d -Dimensional Singularities in \mathbf{R}^n . Let $d \in \{1, \dots, n-1\}$. Is there a system of representation for functions of $L^2(\mathbf{R}^n)$ (e.g., an orthonormal basis) that represents $\tilde{f}_{d,\alpha}$ and similar objects sparsely?

In this note we describe, for each pair (n, k) with $1 \leq k < n$, a construction of a tight frame for $L^2(\mathbf{R}^n)$ that is intended to display the same efficiency of representation of singularities of dimension $k \geq 1$ that wavelets exhibit for singularities of dimension 0. We also sketch arguments suggesting how

to decide when these tight frames might provide the desired efficient representations. Our arguments suggest that these tight frames have the desired sparse coefficient property for $k = n - 1$, for any $n > 1$. They also suggest that, owing to the structure of $G_{n,k}$, the new frames may not have the desired sparse coefficient property in the cases $1 \leq k < n - 1$, $n \geq 3$. In short, it appears that the constructions given here provide efficient representations for $(n - 1)$ -dimensional singularities; and that efficient treatment of lower-dimensional singularities is a topic for further research.

The viewpoint described in this note derives from extensive unpublished work conducted over the last few years by Emmanuel Candès and myself. The article in press (3) and the recent thesis (4) introduced the terminology of ridgelet analysis and the problem of constructing and applying ridgelet frames. Candès' ridgelets are closely related to what we call in the terminology of this paper $(n - 1)$ -plane ridgelet frames; his applications included the potential usefulness of his ridgelet frames for what we call here Problem $(n, n - 1)$. In that pioneer work, the convention was adopted that the term "ridgelet" refers specifically to a ridge function $\psi_{a,b,u}(x) = \psi(au'x + b)a^{1/2}$, where ψ is oscillatory. Here $u \in S^{n-1}$ is a unit vector, and the function $\psi_{a,b,u}$ is constant along $(n - 1)$ -dimensional hyperplanes or "ridges." The phrase "ridgelet frame" referred to frames where the individual elements had this structure, for appropriate (a_n, b_n, u_n) . Working within that constraint, the resulting frames were not tight, the ridgelet primal frame elements had to be accompanied by nonridgelet dual frames, the construction of dual frames was implicit, the properties of the dual frame elements were not available directly, and the primal frame elements were not in $L^2(\mathbf{R}^n)$.

In one of my articles (available at <http://www-stat.stanford.edu/~donoho/Reports/1998/ridge-lin-sing.ps>), I had the idea to broaden the notion of ridgelet so that, rather than imposing the ridge-function form on the elements of an analyzing system, certain localization properties were obeyed in a radial frequency \times angular-frequency domain. Under this broadened notion of ridgelet, it was possible to explicitly construct orthonormal ridgelet bases in the case $k = 1$, $n = 2$ and also to show that the properties of orthonormal ridgelet coefficients can be identified with the properties of tensor wavelet analysis of a fractionally-differentiated Radon transform. This identification was used to show that orthonormal ridgelet coefficients are sparse when analyzing certain smooth objects with linear singularities in \mathbf{R}^2 .

In the present article, we generalize the construction of orthonormal ridgelets in $k = 1$, $n = 2$ to cases of $n > 2$ and k arbitrary. The construction is a generalization of an approach developed in $n = 2$, $k = 1$ because, whereas that case was based on the properties of the Radon transform, the generalization exploits corresponding properties of the k -plane X -ray transform (5). In particular, we rely on Solmon's isometry between X -ray space and real space. In our generalization, we obtain not orthonormal bases but instead tight frames. We also show that properties of the k -plane ridgelet coefficients can be obtained from wavelet analysis of the X -ray transform. As part of this effort, we obtain new systems of representation, such as beamlets in $n = 3$, $k = 1$. We also obtain implications for "classic ridgelets" $k = n - 1$, giving an explicit construction of tight ridgelet frames for all $n \geq 2$, and a method of analysis that suggests that the $(n - 1, n)$ ridgelet coefficients of an object $\tilde{f}_{n-1,\alpha}$ will be sparse.

2. Coordinates for X-Ray Bundles

Let $G_{n,k}$ be the Grassman manifold of unoriented k -planes in \mathbf{R}^n (6-8). Letting $\pi \in G_{n,k}$ be such a k -plane, and π^\perp denote the collection of $x \in \mathbf{R}^n$ orthogonal to π , the X -ray of the

function f in direction π at $x \in \pi^\perp$ is (5)

$$Xf(x, \pi) = \int_\pi f(x + y) dy,$$

where the integral is over $y \in \mathbf{R}^n$ belonging to π . The X -ray transform of f is a function on the fiber bundle $\mathcal{X}_{n,k} = \{(\pi, x), \pi \in G_{n,k}, x \in \pi^\perp\}$, with base space $G_{n,k}$, and each fiber isomorphic to \mathbf{R}^{n-k} . In this section we develop a collection of local coordinates on $\mathcal{X}_{n,k}$.

As a C^∞ manifold, $G_{n,k}$ has an atlas of charts, $\{(N_q, \chi_q), : q = 1, \dots, Q\}$, where each N_q is a neighborhood in $G_{n,k}$ and each χ_q is a diffeomorphism into $\mathbf{R}^{k(n-k)}$. There is a subordinate partition of unity $(w_q : q = 1, \dots, Q)$, where $0 \leq w_q \leq 1$, $\text{supp } w_q \subset N_q$, $\sum_{q=1}^Q w_q(\pi) = 1 \forall \pi$, and each $w_q \in C^\infty$.

In fact, we can be much more specific about the local coordinates for $G_{n,k}$. An element $\pi \in G_{n,k}$ can be associated with the matrices U in the orthogonal group $O(n)$ whose first k -columns give an orthobasis for π . $G_{n,k}$ can therefore be identified with the quotient of $O(n)$ by $O(k) \times O(n - k)$, because any matrix U corresponding to π can be transformed into another such matrix by multiplication with

$$\begin{pmatrix} U_k & 0 \\ 0 & U_{n-k} \end{pmatrix},$$

where U_k is a $k \times k$ orthogonal matrix, etc. Finally, by using the Lie algebra structure of $O(n)$, this quotient group can be associated with the exponentials of skew-symmetric matrices

$$V = \begin{bmatrix} 0 & V_0 \\ -V_0' & 0 \end{bmatrix}, \tag{1}$$

where V_0 is k by $n - k$; the correspondence $\pi \leftrightarrow V$ is C^∞ and one-one on a sufficiently small neighborhood N_q , which in practice may be taken rather large, for example the correspondence is certainly well behaved on all such matrices V of norm $< \pi/4$, say.

Let, for each q , π_q be a point in the interior of N_q , and assign to π_q a matrix $U_q(\pi_q) \in O(n)$ whose first k columns are an orthobasis for π_q . Assuming the neighborhood N_q is appropriately small, we can then define a unique skew-symmetric matrix $V = V(\pi)$ patterned like Eq. 1 such that

$$U_q(\pi) = \exp(V)U_q(\pi_q)$$

gives an orthogonal matrix whose first k columns form an orthobasis for π . Thus, defining a vector $v \in R^{k(n-k)}$ by letting $v = ((V_0)_{1,1}, \dots, (V_0)_{1,n-k}, (V_0)_{2,1}, \dots, (V_0)_{2,n-k}, \dots, (V_0)_{k,n-k})$ be a repackaging of the entries of the submatrix V_0 of V , we may take as definition of our local coordinates

$$v = \chi_q(\pi).$$

This system has the property that $\chi_q(\pi_q) = 0$.

This local coordinate system for the base space $G_{n,k}$ allows us also to define local coordinates on the fiber bundle $\mathcal{X}_{n,k}$. With $U_q(\pi)$ as above, put $U_q(\pi) = [U_q^\parallel(\pi) \mid U_q^\perp(\pi)]$ where U_q^\parallel is n by k and U_q^\perp is n by $n - k$. The columns of $U_q^\parallel(\pi)$ make an orthobasis for π^\perp . Define, for $\pi \in N_q$, corresponding variables

$$u = u_q(\pi, x) = (U_q^\perp(\pi))'x$$

$$v = v_q(\pi) = \text{entries of submatrix } V_0 \text{ of } V.$$

Then on N_q , the correspondence $(\pi, x) \leftrightarrow (v, u)$ is a C^∞ diffeomorphism, and a linear isometry on each $n - k$ -dimensional fiber (π fixed, v fixed). In this way we have constructed a special atlas of charts for $\mathcal{X}_{n,k}$.

3. A Tight Frame on X-Ray Bundles

Define now an L^2 -norm for functions $F(\pi, x) : \mathcal{X}_{n,k} \rightarrow \mathbf{R}$ by

$$\|F\|_{n,k}^2 = \int_{G_{n,k}} \int_{\pi^\perp} |F(\pi, x)|^2 dx d\mu(\pi),$$

where μ is the finite measure on $G_{n,k}$ invariant under an orthogonal transformation of \mathbf{R}^n , normalized so $\mu(G_{n,k}) = |S^{n-1}|/|S^{n-k-1}|$ (compare ref. 5).

The manifold structure underlying $\mathcal{X}_{n,k}$ allows us to write this integration as an integration on Euclidean space $\mathbf{R}^m = \mathbf{R}^{k(n-k)} \times \mathbf{R}^{n-k}$. We may write, for $H(\pi, x)$ a function on $\mathcal{X}_{n,k}$, $H_q(\pi, x) = w_q(\pi)H(\pi, x)$, so that $\text{supp } H_q \subset N_q \times \mathbf{R}^{n-k}$; then

$$\begin{aligned} & \int_{G_{n,k}} \int_{\pi^\perp} H(\pi, x) dx d\mu(\pi) \\ &= \sum_q \int_{N_q} \int_{\pi^\perp} H_q(\pi, x) dx d\mu(\pi) \\ &= \sum_q \int_{\mathbf{R}^{k(n-k)}} \int_{\mathbf{R}^{n-k}} H_q^*(v, u) du J_q(v) dv, \end{aligned}$$

where $H_q^*(v, u) = H_q(\chi_q^{-1}(v), (U_q^\perp(\chi_q^{-1}(v)))'u)$ is the pullback of H_q to Euclidean coordinates, and $J_q(v)$ is the change-of-variables factor guaranteeing

$$\begin{aligned} & \int_{N_q} \int_{\pi^\perp} H_q(\pi, x) dx d\mu = \\ & \int_{\mathbf{R}^{k(n-k)}} \int_{\mathbf{R}^{n-k}} H_q^*(v, u) du J_q(v) dv. \end{aligned} \quad [2]$$

The factor J_q is C^∞ and bounded away from zero on the open set $\chi_q[N_q]$. Here and below, whenever we have a function G defined on a neighborhood of a d -manifold, the the function G^* induced on some subset of \mathbf{R}^d via the local coordinates will be called the *pullback* of G to Euclidean coordinates. In the other direction, whenever we have a function G defined on a subset of \mathbf{R}^d in bijection with a neighborhood in a d -manifold, the function G^* induced on the manifold via the local coordinate system will be called the *pushout* of G from Euclidean space to the manifold. For example, below we will need $J_q^*(\pi) \equiv J_q(\chi_q(\pi))$, the pushout of the change of variables factor from euclidean space to $\mathcal{X}_{n,k}$. This is a C^∞ function on N_q bounded away from 0.

We now use the manifold atlas to construct a wavelet-like frame on $\mathcal{X}_{n,k}$ by a pushout operation from Euclidean space. With $u \in \mathbf{R}^{n-k}$ and $v \in \mathbf{R}^{k(n-k)}$ now variables ranging freely through their respective spaces, define tensor wavelets

$$\psi_\mu(v, u) = \psi_{(j_1, \ell_1, \dots, \ell_{n-k})}^{\delta_1, \dots, \delta_{n-k}}(u) \psi_{(j_2, m_1, \dots, m_M)}^{\varepsilon_1, \dots, \varepsilon_n}(v). \quad [3]$$

Here the terms in the euclidean variable u , $\psi_{(j_1, \ell_1, \dots, \ell_{n-k})}^{\delta_1, \dots, \delta_{n-k}}(u)$, constitute a standard orthonormal basis of Meyer wavelets (9) for \mathbf{R}^{n-k} having “no coarsest level.” The terms in the angular variable v , $\psi_{(j_2, m_1, \dots, m_M)}^{\varepsilon_1, \dots, \varepsilon_n}(v)$, ($M = k(n-k)$) constitute a standard orthonormal basis of Meyer wavelets for $\mathbf{R}^{k(n-k)}$ having a coarsest level j_0 .

In detail, for terms in the Euclidean variable u , the scale index j_1 runs through both positive and negative integers, the ℓ_i index spatial locations, and each δ_i is a binary variable indicating whether the wavelet will be oscillatory in the i th direction or not. Individual terms in this system consist of tensor products of Meyer wavelets $\psi_{j_1, \ell_i}^{\delta_i}(u_i)$ with the constraint that in forming each such product, at least one factor must be oscillatory.

For the angular variable v , the scale index j_2 runs through $j_2 \geq j_0$, where $j_0 \geq 0$. The m_i index spatial locations and each

ε_i is a binary variable indicating again whether the wavelet will be nonoscillatory or oscillatory in that direction. The individual terms again consist of products of $\psi_{j_2, m_i}^{\varepsilon_i}(v_i)$; at scales $j_2 > j_0$, in each such product, at least one oscillatory factor is required. However, for $j_2 = j_0$ only, all products are allowed, including products of all nonoscillatory terms. We let the index $\mu = (j_1; \ell_1, \delta_1, \dots, \ell_{n-k}, \delta_{n-k}; j_2; m_1, \varepsilon_1, \dots, m_{k(n-k)}, \varepsilon_{k(n-k)})$.

For q fixed, pushout these wavelets from Euclidean space to $G_{n,k}$ via

$$\psi_\mu^{*,q}(\pi, x) = \begin{cases} \psi_\mu(v_q(\pi), u_q(\pi, x)); & \pi \in N_q \\ 0 & \text{else} \end{cases}.$$

Note that the behavior of $\psi_\mu^{*,q}$ outside of N_q will not be an issue so that, for example, there is no need for a global definition of coordinates $v(\pi)$. Define frame elements $W_\mu^q(\pi, x)$ by

$$W_\mu^q(\pi, x) = (J_q^*(\pi))^{-1/2} \psi_\mu^{*,q}(\pi, x) w_q^{1/2}(\pi). \quad [4]$$

Note that as w_q is supported in N_q and J_q^* is bounded away from zero on N_q , there is no difficulty with this definition; it defines a C^∞ function supported in N_q . For notational convenience, we will also write W_λ for W_μ^q , where $\lambda = (\mu, q)$.

THEOREM 1. ($W_\lambda(\pi, x)$) $_\lambda$ is a tight frame for $L^2(\mathcal{X}_{n,k})$; if $F : \mathcal{X}_{n,k} \rightarrow \mathbf{R}$ has finite $\|\cdot\|_{n,k}$ -norm, then

$$F = \sum_\lambda [F, W_\lambda] W_\lambda \quad \text{in } L^2(\mathcal{X}_{n,k}) \quad [5]$$

and

$$\|F\|_{n,k}^2 = \sum_\lambda [F, W_\lambda]^2. \quad [6]$$

Proof: Set $F_q(\pi, x) = w_q^{1/2}(\pi)F(\pi, x)$. The transformation $F \mapsto (F_q)_q$ is an isometry from $L^2(\mathcal{X}_{n,k})$ to $\ell^2(L^2(\mathcal{X}_{n,k}))$:

$$\sum_q \|F_q\|_{n,k}^2 = \|F\|_{n,k}^2. \quad [7]$$

[For later use: the transformation $(F_q)_q \mapsto G$ defined by $G(\pi, x) = \sum_q w_q^{1/2}(\pi)F_q(\pi, x)$ is a partial isometry from $\ell^2(L^2(\mathcal{X}_{n,k})) \rightarrow L^2(\mathcal{X}_{n,k})$, with the composition $F \rightarrow (F_q)_q \rightarrow G$ the identity mapping of $L^2(\mathcal{X}_{n,k})$ and the composition $(F_q) \mapsto G \mapsto (G_q)$ an L^2 -projection of $\ell^2(L^2(\mathcal{X}_{n,k}))$.]

The transformation $F_q \mapsto ((F_q, \psi_\lambda^{*,q} J_q^{-1/2}))_\lambda$ is an isometry:

$$\begin{aligned} \|F_q\|_{n,k}^2 &= \int_{N_q} \int_{\pi^\perp} |F_q(\pi, x)|^2 dx d\mu \\ &= \int_{\mathbf{R}^{k(n-k)}} \int_{\mathbf{R}^{n-k}} |F_q^*(v, u)|^2 J_q(v) du dv \\ &= \sum_\mu |\langle F_q^* \cdot J_q^{1/2}, \psi_\mu \rangle|^2. \end{aligned}$$

We remark that, here and below, J_q stands for some C^∞ extension of the function defined on $\chi_q[N_q]$, this extension being bounded and bounded away from zero. Now let G be a function on $N_q \subset \mathcal{X}_{n,k}$ with pullback G^* to $\mathbf{R}^{k(n-k)} \times \mathbf{R}^{(n-k)}$ and let H be a function on $\chi_q[N_q] \subset \mathbf{R}^{k(n-k)} \times \mathbf{R}^{(n-k)}$ with pushout H^* to $\mathcal{X}_{n,k}$; from Eq. 2 $\langle G^* J_q^{1/2}, H \rangle = [G(J_q^*)^{-1/2}, H^*]$. With $G = F_q$ and $H = \psi_\mu$, we get

$$\begin{aligned} \langle F_q^* J_q^{1/2}, \psi_\mu \rangle &= [w_q^{1/2} \cdot F \cdot (J_q^*)^{-1/2}, \psi_\mu^{*,q}] \\ &= [F, W_\lambda] \quad \lambda = (q, \mu). \end{aligned}$$

Hence, for each fixed q ,

$$\|F_q\|_{n,k}^2 = \sum_\mu [F, W_\mu^q]^2. \quad [8]$$

Eq. 6 follows immediately from Eqs. 7 and 8.

To check Eq. 5, note that

$$F_q^* J_q^{1/2} = \sum_{\mu} \langle F_q^* J_q^{1/2}, \psi_{\mu} \rangle \psi_{\mu} \quad \text{in } L^2(\mathbf{R}^{k(n-k)} \times \mathbf{R}^{n-k})$$

because (ψ_{μ}) is a complete orthonormal system. As the extension J_q is bounded and bounded away from zero on all of $\mathbf{R}^{k(n-k)}$, we may write

$$F_q^* = \sum_{\mu} \langle F_q^* J_q^{1/2}, \psi_{\mu} \rangle J_q^{-1/2} \psi_{\mu} \quad \text{in } L^2(\mathbf{R}^{k(n-k)} \times \mathbf{R}^{n-k}).$$

Hence, from $[F, W_{\mu}^q] = \langle F, W_{\mu}^q \rangle$,

$$F_q^* = \sum_{\mu} [F, W_{\mu}^q] J_q^{-1/2} \psi_{\mu}$$

so

$$w_q^{1/2} F_q = \sum_{\mu} [F, W_{\mu}^q] W_{\mu}^q$$

and from $\sum_q w_q^{1/2} F_q = F$ we get Eq. 5. ■

4. Isometry Between X-Ray Space and Real Space

Let $F(\pi, x)$ be a function on $\mathcal{X}_{n,k}$, and let $\tilde{F}(\pi, \xi)$ denote the fiberwise Fourier transform:

$$\tilde{F}(\pi, \xi) = \int_{\pi^{\perp}} F(\pi, x) e^{-i\xi x} dx, \quad \xi \in \pi^{\perp}.$$

We remark that

$$F(\pi, x) = \frac{1}{(2\pi)^k} \int_{\pi^{\perp}} \tilde{F}(\pi, \xi) e^{+i\xi x} d\xi, \quad x \in \pi^{\perp}$$

and that

$$\frac{1}{(2\pi)^k} \int |\tilde{F}(\pi, \xi)|^2 d\xi = \int |F(\pi, x)|^2 dx.$$

It follows that the mapping $\mathcal{F}_{x \rightarrow \xi} F = \tilde{F}$ defines (up to normalization) an isometry from $L^2(\mathcal{X}_{n,k})$ to $L^2(\mathcal{X}_{n,k})$. For such a function $\tilde{F}(\pi, \xi)$ define a new function $\hat{f}(\xi)$ by

$$\hat{f}(\xi) = |\xi|^{-k/2} \tilde{F}(\pi, \xi) \quad [9]$$

(ignoring for the moment the possibility of misbehavior at $\xi = 0$). This gives a mapping from $L^2(\mathbf{R}^n)$ to $L^2(\mathcal{X}_{n,k})$; in fact it is a multiple of an isometry, as

$$\|\tilde{F}\|_{L^2(\mathcal{X}_{n,k})} = \int_{G_{n,k}} \int_{\pi^{\perp}} |\xi|^k |\hat{f}(\xi)|^2 d\xi d\mu = \gamma_{n,k}^0 \|\hat{f}\|_{L^2(\mathbf{R}^n)},$$

for a constant $\gamma_{n,k}^0$ (see ref. 5). Label this ‘‘polar-to-Cartesian’’ operation $C: \hat{f} = C(\tilde{F})$. C is, up to a constant factor, an isometry. Now with \mathcal{F} denoting the standard n -variable Fourier transform, define the linear mapping \mathcal{J} according to

$$\mathcal{J} = \mathcal{F}^{-1} \circ C \circ \mathcal{F}_{x \rightarrow \xi}.$$

At least formally this is a constant multiple of an isometry. The inverse mapping is

$$\mathcal{J}^{-1} = (\mathcal{F}_{x \rightarrow \xi})^{-1} \circ P \circ \mathcal{F},$$

where $P\hat{f} = \tilde{F}$ defines a Cartesian-to-polar mapping from $L^2(\mathbf{R}^n)$ to $L^2(\mathcal{X}_{n,k})$ by

$$\tilde{F}(\pi, \xi) = |\xi|^{k/2} \hat{f}(\xi) \quad \xi \in \pi^{\perp}. \quad [10]$$

An important detail concerns the class of objects f in domain (\mathcal{J}^{-1}) and F in domain (\mathcal{J}) . Suppose that F is a function

on the $\mathcal{X}_{n,k}$ bundle which is *fiberwise highpass*: the fiberwise Fourier transform $\tilde{F}(\pi, \xi) = 0$ for $|\xi| < \Omega_0$ for some $\Omega_0 > 0$. Then certainly Eq. 9 makes sense, and so \mathcal{J} is well defined on F . Similarly, suppose that f is a function on \mathbf{R}^n that is *bandlimited* in the ordinary sense, i.e., so that $\hat{f}(\xi) = 0$ for $|\xi| > \Omega_1$ for some $\Omega_1 < \infty$, then Eq. 10 makes sense, and \mathcal{J}^{-1} is well defined on f .

Now consider the tight frame $(W_{\lambda})_{\lambda}$ for $L^2(\mathcal{X}_{n,k})$. Because in Eq. 3 we chose to use Meyer wavelets in the u -factor, the frame elements W_{λ} obey the fiberwise highpass condition: $\tilde{W}_{\lambda}(\pi, \xi) = 0$ for all $|\xi| < \Omega_{\lambda}$ and all π , for an $\Omega_{\lambda} > 0$. Hence the definition

$$\rho_{\lambda} \equiv \mathcal{J}(W_{\lambda}), \quad \forall \lambda \in \Lambda \quad [11]$$

makes sense and yields an isometric set in $L^2(\mathbf{R}^n)$.

THEOREM 2. (ρ_{λ}) is a tight frame for $L^2(\mathbf{R}^n)$.

Indeed, as (ρ_{λ}) is isometric to (W_{λ}) , it is a tight frame for $\text{span}(\rho_{\lambda})$. It remains to check that this span is all of $L^2(\mathbf{R}^n)$. As bandpass functions are dense in $L^2(\mathbf{R}^n)$, it is enough to check that there are no nonzero bandpass functions orthogonal to every ρ_{λ} ; here by *bandpass* we mean $\hat{f}(\xi) = 0$ for $|\xi| \notin [\Omega_0, \Omega_1]$ for some $0 < \Omega_0 < \Omega_1 < \infty$. The isometry \mathcal{J}^{-1} is well defined on such objects; one sees that $F = \mathcal{J}^{-1}(f)$ defines an isometric function in $L^2(\mathcal{X}_{n,k})$; this F is non-null and the W_{λ} are complete so that $\sum_{\lambda} [F, W_{\lambda}]^2 > 0$. But as the W_{λ} and F are fiberwise highpass, \mathcal{J} is well-defined on every W_{λ} and on F and one sees that term by term we can justify the equality $[F, W_{\lambda}] = \langle \mathcal{J}(F), \mathcal{J}(W_{\lambda}) \rangle = \langle f, \rho_{\lambda} \rangle$, so $\sum_{\lambda} \langle \rho_{\lambda}, f \rangle^2 > 0$. Hence no nontrivial bandpass f can be orthogonal to every ρ_{λ} . ■

5. Examples

5.1. Beamlets. In dimension $n = 3$, let $k = 1$. Then $G_{n,k}$ is the collection of all lines through the origin, commonly denoted P^2 . Given a line π , consider the intersection of that line with S^2 , the unit sphere in 3-space. This consists of two antipodal points $\{p, -p\}$. We may identify $\pi \sim \{p, -p\}$, showing that S^2 gives a double covering of $G_{n,k}$.

We work concretely now on the sphere, keeping in mind the picture that $\pi \sim \{p, -p\}$. For $q = 1, 2, 3$, consider the q th polar cap Γ_q^1 the region of $p \in S^2$ where $p_q = \max_i |p_i|$, and the slightly larger cap Γ_q^0 where $p_q \geq \frac{1}{2} \cdot \max_i |p_i|$. We can define windows ν_q so that $\nu_q(p) = 1$ on Γ_q^1 , so that $\nu_q(p) = 0$ outside Γ_q^0 . We can create a partition of unity on P^2 by the recipe $\sigma_q(\pi) = \nu_q(\bar{p}_q(\pi))$, where $\bar{p}_q(\pi)$ means ‘‘the element p of the antipodal pair associated with π that lies closest to the q th pole.’’ Defining then $w_q(\pi) = \sigma_q(\pi) / \sum_q \sigma_q(\pi)$, we have a smooth partition of unity isolating these three regions, i.e., smooth functions $w_q(\pi)$ so that $0 \leq w_q(\pi) \leq 1$, $\sum_q w_q(\pi) \equiv 1, \forall \pi \in P^2$, and so that $w_q = 1$ at lines coming near the q th pole, reaching zero at lines near the corresponding equator.

For the q th local coordinate system on $G_{3,1}$, let U_q be a permutation matrix whose first column has a one in the q th position, and let $U_q(v) = \exp(V)U_q$. The first column of $U_q(v)$ defines a smooth bijection from Euclidean coordinates v onto the q th polar cap. This of course defines a point $p = \bar{p}_q(\pi)$ in correspondence with a line π which is the same point as that defined by the rule of the previous paragraph. The second and third columns of $U_q(v)$ define coordinates for the orthocomplement π^{\perp} . Letting u be the coordinates of π^{\perp} in this coordinate system, we can obtain a tight frame for $\mathcal{X}_{3,1}$ using Eq. 4. We then realize a tight frame for $L^2(R^3)$ by applying the X -ray isometry. To make this more explicit, we have the frequency-domain formula

$$\hat{\rho}_{\lambda}(\xi) = \phi_{q,\mu_0}(\pi) \hat{\psi}_{\mu_1}^+((U^{\perp}(\pi))' \xi), \quad \xi \in \pi^{\perp}.$$

Here the factor $\phi_{q,\mu_0}(\pi) \equiv w_q^{1/2}(\pi)(J_q^*)^{-1/2}(\pi)\psi_{\mu_0}^{(0)}(v_q(\pi))$ is an element of a tight frame on P^2 , $\psi_{\mu_0}^{(0)}(v)$ being a tensor wavelet on \mathbf{R}^2 ; and, with $\psi_{\mu_1}^{(1)}(u)$ a tensor wavelet on R^2 , $\psi_{\mu_1}^+$ is a wavelet-like function on \mathbf{R}^2 defined in the Fourier domain by $\hat{\psi}_{\mu_1}^+(\omega) = |\omega|^{1/2}\hat{\psi}_{\mu_1}^{(1)}(\omega)$. Thus $\lambda = (q, \mu_0, \mu_1)$ groups together the index of the polar cap, the index of the wavelet expansion on the local coordinates for that cap, and the index of the wavelet expansion on the euclidean coordinates for π^\perp .

The family of induced functions ρ_λ concentrates “near” beams, hence they may be called *beamlets*.

5.2. Classic Ridgelets. Consider again the case $n = 3$, now with $k = 2 = n - 1$, which is the setting for “classic ridgelet analysis” of ref. 3. Now 2-plane ridgelets should concentrate near planes in \mathbf{R}^3 , so they might also be called “platelets.” Each plane in $G_{3,2}$ may be identified with its orthocomplement—a line in $G_{3,1}$; i.e., we are indexing position by using π^\perp rather than π . Once again, the sphere S^2 provides a double cover and a concrete manifold to work with. We may use the same polar cap partition of unity ($w_q : q = 1, 2, 3$) as with beamlets; as π^\perp is one-dimensional, the implicit function $p = \bar{p}_q(\pi^\perp)$ defines a basis for π^\perp . From the tight frame formula 4 we can specialize to the case at hand, and get a frequency-domain formula for $\hat{\rho}_\lambda$. For λ fixed, and hence q fixed, we may write $\xi \in \pi^\perp$ as $\xi = p \cdot \omega$ with $p = \bar{p}_q(\pi^\perp)$, and $\omega \in \mathbf{R}$, and then

$$\hat{\rho}_\lambda(\xi) = \phi_{q,\mu_0}(\pi^\perp) \cdot \hat{\psi}_{\mu_1}^+(\omega),$$

where, as in the beamlet case, $\phi_{q,\mu_0}(\pi^\perp) = w_q^{1/2}(\pi^\perp) \cdot \psi_{\mu_0}^{(0)}(v_q(\pi^\perp))(J_q^*)^{-1/2}(\pi^\perp)$ is an element of a tight frame on P^2 [$\psi_{\mu_0}^{(0)}(v)$ being a tensor wavelet on \mathbf{R}^2], and we define $\hat{\psi}_{\mu_1}^+$ in terms of $\psi_{\mu_1}^{(1)}$, a wavelet on \mathbf{R}^1 , by the frequency-domain formula $\hat{\psi}_{\mu_1}^+(\omega) = \hat{\psi}_{\mu_1}^{(1)}(\omega) \cdot |\omega|$. So the frequency-domain structure of the tight frame is of a “spherical wavelet” times the Fourier transform of a “wavelet on the line.”

6. Interpretation in X-Ray Space

Again in the setting of general n and k , let $F(\pi, x)$ be a “sufficiently nice” function on $\mathcal{X}_{n,k}$ and let X^* be the formal adjoint of the X -ray transform

$$(X^*F)(x) = \gamma_{n,k} \int_{G_{n,k}} F(\pi, \pi^\perp(x)) d\mu(\pi),$$

where $\pi^\perp(x)$ denotes orthogonal projection of x onto π^\perp . The constant $\gamma_{n,k}$ is chosen so that on sufficiently “nice” pairs F, g with $F \in L^2(\mathcal{X}_{n,k})$ and $g \in L^2(\mathbf{R}^n)$,

$$\langle X^*F, g \rangle = [F, Xg]. \tag{12}$$

Here “sufficiently nice” means, in addition to typical localization and smoothness conditions, that F is fiberwise highpass.

Given now a function $F(\pi, x)$ on $\mathcal{X}_{n,k}$, let D^+, D^- be operators defined fiberwise on such functions as follows. With again $\tilde{F}(\pi, \xi)$ the fiberwise Fourier transform, set

$$D^+F(\pi, x) = \frac{1}{(2\pi)^k} \int_{\pi^\perp} e^{ix\xi} |\xi|^{k/2} \tilde{F}(\pi, \xi) d\xi, \quad x \in \pi^\perp$$

and

$$D^-F(\pi, x) = \frac{1}{(2\pi)^k} \int_{\pi^\perp} e^{ix\xi} |\xi|^{-k/2} \tilde{F}(\pi, \xi) d\xi, \quad x \in \pi^\perp.$$

Assuming that F is fiberwise bandpass, these expressions make sense rigorously. Now the W_λ are indeed bandpass on each fiber, so D^\pm are well defined on such functions; set

$$\tau_\lambda = D^+W_\lambda, \quad \sigma_\lambda = D^-W_\lambda;$$

these are a roughening and smoothing of W_λ , respectively. The adjoint equation 12 gives immediately the following theorem:

THEOREM 3. *We have the operator biorthogonality relations*

$$\begin{aligned} \rho_\lambda &= X^*(\tau_\lambda) & \sigma_\lambda &= X(\rho_\lambda), & \lambda \in \Lambda \\ \langle \rho_\lambda, \rho_{\lambda'} \rangle &= [\sigma_\lambda, \tau_{\lambda'}] & &= [W_\lambda, W_{\lambda'}], & \lambda, \lambda' \in \Lambda. \end{aligned}$$

If f is a finite sum of ρ_λ 's then

$$f = \sum_\lambda [\tau_\lambda, Xf] \rho_\lambda$$

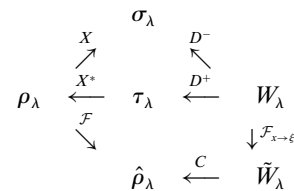
$$\|f\|_{L^2(\mathbf{R}^n)}^2 = \sum_\lambda [\tau_\lambda, Xf]^2.$$

In short, the coefficients of the ρ_λ expansion can be read off from the τ_λ analysis of the X -ray transform. As $\tau_\lambda = D^+W_\lambda$, and D^+ is self-adjoint on the appropriate domain, we may also write for such f

$$f = \sum_\lambda [W_\lambda, D^+Xf] \rho_\lambda$$

$$\|f\|_{L^2(\mathbf{R}^n)}^2 = \sum_\lambda [W_\lambda, D^+Xf]^2.$$

So the coefficients can be read off from a fiberwise differentiated version of the X -ray transform. We have the following commuting diagram:



7. Analysis of a k-Dimensional Singularity

We now discuss at an informal level the possible effectiveness of k -plane ridgelets at representing k -dimensional singularities. Consider the object

$$f(x) = s(x_1, \dots, x_{n-k})w(x_{n-k+1}, \dots, x_n),$$

where, say, s contains a singularity at $x_1 = \dots = x_{n-k} = 0$ (say of the form $s \sim (x_1^2 + \dots + x_{n-k}^2)^{-\alpha/2}$ as $(x_1, \dots, x_d) \rightarrow 0$) and w is C^∞ of compact support. Section 6 showed that k -plane ridgelet analysis of f is the same as wavelet analysis of the fiberwise differentiated $Xf(\pi, x)$. And, more or less, wavelet analysis gives sparse coefficients when an object is smooth away from a point singularity. Hence we are interested in the following question: when is D^+Xf smooth away from a point singularity at $\pi = \pi_0, x = 0$? If in fact D^+Xf has this property, then we can expect, based on known principles, for k -plane ridgelet analysis of f to give sparse coefficients, whereas if D^+Xf does not have this property, sparsity of the coefficients will not follow from known principles.

We consider first the question of smoothness from a frequency-domain perspective. For the n -dimensional Fourier transform of f we have

$$\hat{f}(\xi) = \hat{s}(\xi^0)\hat{w}(\xi^1) \quad \xi = (\xi^0, \xi^1),$$

where $\xi^0 = (\xi_1, \dots, \xi_{n-k}), \xi^1 = (\xi_{n-k+1}, \dots, \xi_n)$. Now as s has a singularity at the origin, \hat{s} is not of rapid decay, but \hat{w} is of rapid decay. Defining coordinate projections $\xi_0(\xi) = \xi^0$ and $\xi_1(\xi) = \xi^1$, we obtain for the fiberwise Fourier transform of $F = Xf$

$$\tilde{F}(\pi, \xi) = \hat{s}(\xi_0(\xi))\hat{w}(\xi_1(\xi)) \quad \xi \in \pi^\perp. \tag{13}$$

We are interested in the question of whether this expression is of rapid decay as $|\xi| \rightarrow \infty$ for $\pi \neq \pi_0$. Letting π_0^\perp denote the

subspace $\xi_1 = 0$, it is clear that the expression will not typically be of rapid decay when $\pi = \pi_0$. Indeed, if $\int w \neq 0$, then for $\xi \in \pi_0^\perp$, $\hat{w}(\xi_1(\xi)) = \hat{w}(0)$, so $\tilde{F}(\pi_0, \xi) = \hat{s}(\xi^0)\hat{w}(0)$, which fails to be of rapid decay, because \hat{s} is not of rapid decay. However, expression 13 will be of rapid decay in ξ provided that π^\perp contains no nontrivial subspace lying entirely also in π_0^\perp , i.e., provided the minimum angle between the subspaces π^\perp and π_0^\perp is strictly positive. For, under this minimum angle condition, if we tend to ∞ along any line in π^\perp , the factor $|\xi_1|$ tends to ∞ ; as \hat{w} is of rapid decay, this forces $\hat{w}(\xi_1(\xi))$ to be of rapid decay on $\xi \in \pi^\perp$; by 13, this means $\tilde{F}(\pi, \xi)$ is of rapid decay as $|\xi| \rightarrow \infty$ in π^\perp .

Slightly more elaborately, we may see the same thing in the X-ray domain rather than the Fourier domain. For the X-ray transform of f we have the fiberwise convolution

$$F(\pi, x) = \int_{\pi^\perp} s(Ax)|A|^{-1} \cdot w(B(x-x'))|B|^{-1}dx', \quad x \in \pi^\perp, \quad [14]$$

where the convolution takes place in the fiber π^\perp , and the mappings $A = A(\pi, \pi_0)$, $B = B(\pi, \pi_0)$, satisfy

$$\begin{aligned} A|_{\pi^\perp} &\rightarrow A_0|_{\pi_0^\perp} \quad \text{as } \pi \rightarrow \pi_0 \\ B|_{\pi^\perp} &\rightarrow 0 \quad \text{as } \pi \rightarrow \pi_0. \end{aligned}$$

Here A_0 selects the first $n - k$ components in the standard basis. Also, $|A|^{-1}$ refers to the determinant of the matrix of $\dim(\pi^\perp)$ by $\dim(\pi^\perp)$ representing $A : \pi^\perp \rightarrow \mathbf{R}^{n-k}$; $|B|^{-1}$ refers similarly to the mapping $B : \pi^\perp \rightarrow \mathbf{R}^k$. The mapping B may not be of full rank if π^\perp contains a line in π_0^\perp , in which case we interpret the degenerate expression $w(B(x-x'))|B|^{-1}$ as a suitable generalized function, obtained by the obvious limiting process through full rank matrices $\tilde{B} \rightarrow B$.

Under the minimum angle condition, B will be of full rank, $w(B(x-x'))|B|^{-1}$ will be nondegenerate and C^∞ , and we can see that $F(\pi, x)$ is C^∞ on that fiber, because we may push all differentiations in x over to this nondegenerate w factor. If the minimum angle condition is violated, B is degenerate, and $w(B(x-x'))|B|^{-1}$ behaves as a generalized function in certain slices, so we lose the ability to push all differentiations over to the w factor. As a result, F is not C^∞ on such a fiber.

Now if $k = n - 1$, the question of comparing π^\perp and π_0^\perp is merely a question of comparing one-dimensional subspaces. In such a case, if $\pi \neq \pi_0$ then the minimum angle condition is met. Hence, if $k = n - 1$, $F(\pi, x)$ is C^∞ whenever $\pi \neq \pi_0$; it has a pure point singularity at $\pi = \pi_0, x = 0$.

Hence, for $k = n - 1$, k -plane ridgelet analysis amounts to a wavelet analysis of a function with a point singularity. In dimension $n = 2, k = 1$, additional calculations show that wavelet coefficients of such a function will be sparse. We expect similar findings here; so it seems likely that $n - 1$ -plane ridgelets solve Problem($n, n - 1$) of the introduction for all $n > 1$.

If $1 \leq k < n - 1$ and $n \geq 3$, in comparing π with a nearby π_0 we may very easily have $\pi^\perp \neq \pi_0^\perp$ while at the same time the minimum angle between the two subspaces is zero. In that case, $\tilde{F}(\pi, \xi)$ will not be of rapid decay in certain directions in the fiber, and $F(\pi, x)$ will not be C^∞ in the fiber. Hence the function $F(\pi, x)$ will fail to have C^∞ behavior in x not only at

$\pi = \pi_0$, but also along certain sections (curves or surfaces) in $G_{n,k}$ originating from π_0 .

Although wavelet analysis can yield sparse coefficients for point singularities, we know of no precedent to suggest that it can yield sparse coefficients on functions with nonsmooth directions in certain sections originating at a point. This is not to say that sparsity is ruled out, but that if it holds, it must involve rather delicate analysis.

8. Discussion

We have constructed tight frames of k -plane ridgelets, motivated by the possibility that they might allow efficient representation of d -dimensional singularities in \mathbf{R}^n . The analytic machinery of Sections 6 and 7 suggests that we are in a position to now make progress on the following questions:

Q1. Can the tight frames of $(n - 1)$ -plane ridgelets developed here indeed give sparse representations of functions $\tilde{f}_{n-1,\alpha}$?

Q2. Can the tight frames of k -plane ridgelets developed here give sparse representations of $\tilde{f}_{k,\alpha}$ when $1 \leq k < n - 1$?

Q3. If the answer to Q2 is negative, will some other construction be found to solve Problem(n,d) for $d < n - 1$?

In effect, we know the full story on questions Q1-Q3 for $n = 2$, where really only Q1 makes sense; from additional calculations, the answer to Q1 is affirmative in dimension $n = 2$. Going next to $n = 3$, all three questions make sense, and are both interesting and challenging. The analytic machinery of Sections 6 and 7, and the experience with an analogous machinery in dimension $n = 2$ suggests that the answer to Q1 is probably yes, and the answer to Q2 may well be no.

Affirmative answers to Q1 and Q2 are expected to have applications in a range of disciplines. Already in the case $n = 3$, the names "beamlets" and "platelets" suggest potential areas of application.

This work is dedicated to my father, Dr. Paul L. Donoho, physicist. It is a pleasure to acknowledge many helpful comments from Ingrid Daubechies on an earlier draft, wide-ranging discussions on ridgelets with Emmanuel Candès, and helpful information about Lie Algebra coordinates for Grassman manifolds provided by Arieh Iserles and Stanislaw Szarek. This work was supported by National Science Foundation Grant DMS-95-05151 and by Air Force Office of Scientific Research Grant MURI95-F49620-96-1-0028.

1. Donoho, D. L., Vetterli, M., Daubechies, I. & DeVore, R. A. (1998) *IEEE Trans. Info. Thry.* **44**, 2435-2476.
2. Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. & Picard, D. (1995) *J. R. Stat. Soc. B* **57**, 301-369.
3. Candès, E. J. (1999) *Appl. Comput. Harmon. Anal.*, in press.
4. Candès, E. J. (1998) Ph.D. thesis (Stanford University, Stanford, CA).
5. Solmon, D. C. (1976) *J. Math. Anal. Appl.* **56**, 61-83.
6. Helgason, S. (1978) *Differential Geometry, Lie Groups, and Symmetric Spaces* (Academic, New York).
7. Spivak, M. (1979) *A Comprehensive Introduction to Differential Geometry* (Publish or Perish, Inc., Wilmington, DE), Vol. 5, 2nd Ed.
8. Boothby, W. M. (1986) *An Introduction to Differentiable Manifolds and Riemannian Geometry* (Academic, Orlando, FL), 2nd Ed.
9. Lemarié, P. G. & Meyer, Y. (1986) *Revista Matematica Iberoamericana* **2**, 1-18.