ON SINGULAR VALUE DECOMPOSITIONS FOR THE RADON TRANSFORM
AND SMOOTHNESS CLASSES OF FUNCTIONS

BY

I. M. JOHNSTONE

TECHNICAL REPORT NO. 310
JANUARY 1989

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT DMS96-00235

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ON SINGULAR VALUE DECOMPOSITIONS FOR THE RADON TRANSFORM
AND SMOOTHNESS CLASSES OF FUNCTIONS

BY

I. M. JOHNSTONE

TECHNICAL REPORT NO. 310
JANUARY 1989

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT DMS86-00235

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
On Singular Value Decompositions for the Radon Transform and Smoothness Classes of Functions

I. M. Johnstone*
Stanford University

Abstract

An exposition is given of the singular value decomposition of the Radon transform for a class of rotationally invariant measures on $\mathbb{R}^2$. The decomposition leads to ellipsoid conditions for smoothness of a function based on ellipsoid constraints on its Fourier coefficients. The report provides reference material for two recent applications in projection approximation for regression and in positron emission tomography.

* Prepared with the support of NSF (MCS 84-51750, 86-00235) and SERC at Stanford University and the University of Bath.
1. Introduction

These notes describe the singular value decomposition (SVD) of the (normalized) Radon transform (Marr (1974), Deans (1983, §7.6) and the original sources quoted therein) with respect to a special one parameter family of measures on the unit disk in $\mathbb{R}^2$. This description is given in a probabilistic notation needed for two distinct, but related applications. (Johnstone and Silverman, 1988 and Donoho and Johnstone, 1986, 1988).

Firstly, in the context of positron emission tomography (PET), we describe how rates of decay of generalized Fourier coefficients $\{f_\nu\}$ of a function $f$ (expressed in ellipsoid conditions) correspond to Sobolev-type smoothness conditions on $f$. It will be seen, however, that these smoothness conditions involve a modified weight function, still contained within the original one-parameter family.

The second application arises in studying how well a function of two variables is approximated by sums of ‘plane waves’ or ‘ridge’ functions. Donoho and Johnstone (1989) study the case of a Gaussian weight measure, and these notes collect the tools needed to look at the unexpected changes that occur in these results as one varies the underlying measure.

The development of the SVD of our “conditional expectation” version of the Radon transform follows that of Davison and Grunbaum (1981) and Davison (1983) except in matters of notation. It is recapitulated here for reference since careful (and tedious!) bookkeeping with the constants is needed for our interpretation of smoothness conditions in terms of ellipsoid conditions on generalized Fourier coefficients. In addition to the SVD, the main conclusions are formula (13), which computes the action of differentiation on the singular functions, and the Theorem, which expresses Sobolev type smoothness properties in terms of an ellipsoid condition on the decay of generalized Fourier coefficients.

2. The condition expectation transform and applications

Define one parameter families of probability measures on $B = \{x \in \mathbb{R}^2 : |x| \leq 1\}$ and $D = \{(s, \phi) \in [-1,1] \times [0,2\pi]\}$ by

\[ d\mu_\alpha = \alpha(1 - |x|^2)^{\alpha-1}d\mu, \quad d\lambda_\alpha = 2^{2\alpha-2}B(\alpha,\alpha)(1-s^2)^{\alpha-1}d\lambda, \quad \alpha > 0. \]  

Here $d\mu_1(x) = d\mu(x) = \pi^{-1}dx$, $d\lambda_1(s,\phi) = d\lambda(s,\phi) = \pi^{-2}(1-s^2)^{1/2}ds\,d\phi$ and $B(\alpha,\alpha) = \Gamma^2(\alpha)/\Gamma(2\alpha)$.

Let $X^\alpha$ denote a random vector with distribution given by $\mu_\alpha$. The conditional expectation transform $P^\alpha: L^2(B,\mu_\alpha) \to L^2(D,\lambda_\alpha)$ is defined by

\[ (P^\alpha f)(s,\phi) = E[f(X^\alpha)|u_\phi \cdot X^\alpha = s] \quad u_\phi = (\cos \phi, \sin \phi)^t \]
The measures \( \{ \lambda_\alpha \} \) and transforms \( \{ P^\alpha \} \) arise in at least two distinct, but related ways. First, in the PET context of (JS, Section 2), a point \( X^\alpha \) is chosen according to \( \mu_\alpha \), and then a random direction \( \Psi \), uniform on \([0, 2\pi]\). The resulting triple \((X_1, X_2, \Psi)\) is reparametrized as \((S, \Phi, T)\), where \( \Phi \) is the angle to horizontal and \( S \) the perpendicular distance to the origin of the line \( \ell \) through \((X_1, X_2)\) in direction \( \Psi \), and \( T \) is the (signed) distance along \( \ell \) of \((X_1, X_2)\) from the closest point on \( \ell \) to 0. This parametrization is one-to-two, since \((S, \Phi, T)\) and \((S, \Phi + \pi, T)\) correspond to the same point \((X_1, X_2, \Phi)\). The measure \( \lambda_\alpha \) is the marginal distribution of \((S, \Phi)\) when \( X^\alpha \) is distributed according to \( \mu_\alpha \). (We deal with the indeterminacy by requiring that \( \lambda_\alpha \) be invariant under the mapping \((s, \phi) \mapsto (-s, \phi + \pi)\)).

If the emission density of \((X_1, X_2)\) is given by \( f(x) d\mu_\alpha(x) \), the corresponding density of \((s, \phi)\) may be written as \( g(s, \phi) d\lambda_\alpha(s, \phi) \), and a short calculation shows that

\[
g(s, \phi) = \frac{\int f(s \cos \phi - t \sin \phi, s \sin \phi + t \cos \phi)(1 - s^2 - t^2)^{\alpha-1} dt}{\int(1 - s^2 - t^2)^{\alpha-1} dt}.
\]

In terms of the original coordinates \( x_1 = s \cos \phi - t \sin \phi, x_2 = s \sin \phi + t \cos \phi \), this becomes

\[
g(s, \phi) = E[f(X^\alpha)|u_\phi \cdot X^\alpha = s] = (P^\alpha f)(s, \phi).
\]

The second setting derives from projection pursuit regression (Friedman and Stuetzle (1981), Huber (1985), Donoho and Johnstone (1988, 1989)). Suppose that \( n \) observations are taken from the model \( Y = f(X) + \epsilon \), where \( X \sim \mu_\alpha \) and \( \epsilon \) is a zero-mean noise variable independent of \( X \). If one looks at a projection \((u_\phi \cdot X_1, Y_i)\) \((i = 1, \cdots, n)\) of the data in direction \( u_\phi = (\cos \phi, \sin \phi) \), the mean of \( Y \) in the projection is given by

\[
g(s, \phi) = E[Y|u_\phi \cdot X^\alpha = s] = E[f(X^\alpha)|u_\phi \cdot X^\alpha = s] = (P^\alpha f)(s, \phi).
\]

From this perspective, \( \lambda_\alpha \) arises as the product of uniform measure on \( \phi \) with the marginal distribution (corresponding to \( u_\phi \cdot X^\alpha \)) of the radially symmetric, bivariate measure \( \mu_\alpha : (g, g') \lambda_\alpha = \frac{1}{2\pi} \int_0^{2\pi} d\phi \int_0^{\infty} \mu_\alpha (X^\alpha) g'(u_\phi \cdot X) \). One can also check that

\[
\frac{d\lambda_\alpha}{d\lambda}(s, \phi) = E \left[ \frac{d\mu_\alpha}{d\mu}(X^\alpha)|u_\phi \cdot X^\alpha = s \right].
\]

The family \( \{ \mu_\alpha \} \) contains uniform measure \((\alpha = 1)\) and Gaussian measure (by rescaling the radius of \( B \) to \( R = \sqrt{2\alpha} \) as \( \alpha \to \infty \)). It is the only compactly supported rotationally invariant family of measures with the property that spaces of polynomials are invariant under the conditional expectation operation \( P^\alpha \) (cf. Davison and Grunbaum, 1981, §4).
3. The adjoint operator to $P$

Although the conditional expectation operators $P^\alpha$ depend on $\alpha$, their adjoints do not. We begin by considering the sections $(P^\alpha_\phi f)(s) = (P^\alpha f)(s, \phi)$ and note that their adjoints are simply “backprojection”:

$P^\alpha_\phi : L^2(B, \mu_\alpha) \rightarrow L^2(\pi_1 D, \pi_1 \lambda_\alpha)$ is linear, with norm 1 and adjoint $(P^*_\phi h)(x) = h(u_\phi \cdot x)$ (here, $\pi_1$ denotes projection on the first coordinate).

[(Throughout, text enclosed in square brackets $[\cdots]$ contains details of proofs of preceding statements.) Let $E_\alpha$ denote expectation with respect to $\mu_\alpha$. The adjoint property follows from the identities

$$E_\alpha P^\alpha_\phi f(u_\phi \cdot X) g(u_\phi \cdot X) = E_\alpha [E[f(X)|u_\phi \cdot X] g(u_\phi \cdot X)]$$

$$= E_\alpha f(X) (P^*_\phi g)(X).$$

The unit norm is a consequence of the Cauchy-Schwarz inequality $E_\alpha E^2_\phi(f|u_\phi \cdot X) \leq E_\alpha f^2(X)$, in which equality holds if $f \equiv 1$.]

Simultaneous diagonalization of $P_\phi P^*_\phi$ provides the key to the singular value decomposition of $P$. The Gegenbauer polynomials of index $\alpha$, $\gamma_m(s) = C^\alpha_m(s)$, $m = 0, 1, 2, \cdots$, form a complete orthogonal basis for $L^2(\pi_1 D, \pi_1 \lambda_\alpha)$ (when normalized, say, as in Gradshteyn-Rhyzik or Magnus et al. (1966)). Furthermore

$$P_\phi P^*_\phi \gamma_m = A_m(\phi - \theta) \gamma_m = \frac{C^\alpha_m(\cos(\phi - \theta))}{C^\alpha_m(1)} \gamma_m. \quad (3)$$

In particular, for $\alpha = 1$, $C^1_m(\cos \theta) = U_m(\cos \theta) = \sin(m + 1)\theta/\sin \theta$ is the $m$th Chebyshev polynomial of the second kind.

[Formula (3) is the main theorem (3.1) of Grunbaum-Davison (1981). Note that we use different definitions of $P_\phi$ (and hence $P^*_\phi$). To connect to their proof, note that

$$(P_\phi P^*_\phi C^\alpha_m)(s) = E_\alpha \left[ C^\alpha_m \left( u^t_{\theta - \phi} \begin{pmatrix} S \\ T \end{pmatrix} \right) \bigg| S = s \right] = \frac{\int C^\alpha_m(s \cos(\theta - \phi) + t \sin(\theta - \phi))(1 - s^2 - t^2)^{\alpha-1} dt}{\int(1 - s^2 - t^2)^{\alpha-1} dt}.$$

The adjoint $P^*$ of $P$ is obtained by averaging $P^*_\phi$ over $\phi$: $P : L^2(B, \mu_\alpha) \rightarrow L^2(D, \lambda_\alpha)$ is linear, with norm 1 and adjoint

$$(P^* g)(x) = \frac{1}{2\pi} \int_0^{2\pi} d\phi \ g(x \cdot u_\phi, \phi) \quad (4)$$

[To verify that (4) gives $P^*$, simply unwind definitions:

$$(f, P^* g)_{\mu_\alpha} = E_\alpha f(X) \frac{1}{2\pi} \int_0^{2\pi} g(u_\phi \cdot X) d\phi = \frac{1}{2\pi} \int_0^{2\pi} E_\alpha g(u_\phi \cdot X) E[f(X)|u_\phi \cdot X] = (P f, g)_{\lambda_\alpha}. \quad (5)$$]
The proof that $P$ has norm 1 is very similar to that for $P_\phi$:

$$\|Pf\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} d\phi \|P_\phi f\|_2^2, \lambda_\alpha \leq \|f\|^2,$$

with equality holding when $f \equiv 1.$


The spectrum of $PP^*$ is obtained from the decomposition of $L^2(D, \lambda_\alpha) = \bigoplus_{m=0}^{\infty} V_m$ into $PP^*$-invariant subspaces $V_m = \text{span}\{\gamma_m(s)v(\phi) : v \in L^2[0, 2\pi]\}$, where

$$(PP^* \gamma_m v)(s, \phi) = \gamma_m(s) \cdot \frac{1}{2\pi} \int_0^{2\pi} A_m(\phi - \theta)v(\theta) d\theta.$$  \hspace{1cm} (5)

Eigenfunctions of $PP^*$ are proportional to $e^{i\ell\phi}\gamma_m(s)$, while eigenvalues $b_{\ell m}^2$ of $PP^*$ are non-zero on $N = (\ell, m) \in \mathbb{Z}^2 : \ell \in \{m, m-2, \ldots, -m\}$ and $m = 0, 1, 2, \ldots$. In the $(j, k)$ index system defined by $j + k = m$, $j - k = \ell$, these eigenvalues are found (after calculation) to be

$$b_{\ell m}^2 = b_{jk}^2 \left( \frac{\alpha_j(\alpha_k)}{(2\alpha)_m} \right)_{j \geq 0, \ k \geq 0}$$ \hspace{1cm} (6)

where $(\alpha)_j = \Gamma(\alpha + j)/\Gamma(\alpha) = \alpha(\alpha + 1) \cdots (\alpha + j - 1)$.

In particular, when $\alpha = 1$, $PP^*$ has eigenfunctions $e^{i\ell\phi}\gamma_m(s)$ and eigenvalues $b_{\ell m}^2 = (1 + m)^{-1}$ for $\ell = m, m - 2, \ldots, -m$.

[The action (5) of $PP^*$ on $V_m$ follows from the diagonalization (3):

$$(PP^* \gamma_m v)(s, \phi) = P \left[ \frac{1}{2\pi} \int_0^{2\pi} d\theta \gamma_m(u_\theta \cdot x)v(\theta) \right](s, \phi)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} d\theta (P_\phi P_\theta^* \gamma_m)(s) v(\theta)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} d\theta A_m(\phi - \theta) \gamma(s) v(\theta).$$

Equation (5) reduces verification of the eigenfunctions of $PP^*$ to the change of variables $\theta = \phi + \phi'$

$$\frac{1}{2\pi} \int_0^{2\pi} A_m(\phi - \theta)e^{i\ell\theta} d\theta = \frac{1}{2\pi} \int_0^{2\pi} A_m(-\phi')e^{i\phi} e^{i\ell\phi'} d\phi'$$

$$= b_{\ell m}^2 e^{i\ell\phi}.$$

To calculate the eigenvalues, employ (3) for the (even) function $A_m$, together with the addition formula for Gegenbauer polynomials [e.g. Gradshteyn and Ryzhik 8.934.2] and the identity $C_m^\alpha(1) = \Gamma(2\alpha + 1)/\Gamma(2\alpha)m!$:

$$b_{\ell m}^2 = \frac{1}{2\pi} \frac{1}{C_m^\alpha(1)} \int_0^{2\pi} C_m^\alpha(\cos \theta)(\cos \ell \theta + i \sin \ell \theta)$$

$$= \sum_{j, k \geq 0} \frac{\Gamma(2\alpha)}{\Gamma(2\alpha + m)} \frac{\Gamma(\alpha + k)}{\Gamma(\alpha + j)} \frac{1}{\Gamma(\alpha)k!} \frac{1}{\Gamma(\alpha)j!} \frac{1}{2\pi} \int_0^{2\pi} \cos(k - j)\theta \cos \ell \theta d\theta.$$
The integral vanishes unless \( \ell = \pm (k - j) \), and this occurs for \((j, k)\) in the range of the summation exactly when \( \ell \in \{m, m - 2, \ldots, -m\} \). For such \( \ell \) there are two (equal) terms in the sum when \( \ell \neq 0 \), and one if \( \ell = 0 \). Hence

\[
b_{\ell m}^2 = \frac{\Gamma(2\alpha)m!}{\Gamma(2\alpha + m)} \frac{\Gamma(\alpha + k) \Gamma(\alpha + j)}{\Gamma(\alpha)k! \Gamma(\alpha)j!}, \quad j + k = m, \quad j - k = \ell \in \{m, m - 2, \ldots, -m\}
\]

which yields the formula in (6).]

The singular value decomposition (SVD) of \( P : L^2(B, \mu_\alpha) \to L^2(D, \lambda_\alpha) \) is now built in a standard way. Choose constants \((c_m^2)^2 = \Gamma(2\alpha)(m + \alpha)!/[\alpha \Gamma(m + 2\alpha)]\), from GR 7.313.2 so that \( \psi_\nu = \psi_{\ell m}^\alpha(s, \phi) = c_m^\alpha e^{i\ell \theta} C_m^\alpha(s), \nu \in N \), is an orthogonal basis for the range of \( P \) in \( L^2(D, \lambda_\alpha) \). 

\( P^* \) pulls \( \{\psi_\nu\} \) back to an orthonormal basis \( \{\phi_\nu\} \) for \( L^2(B, \mu_\alpha) \):

\[
\phi_\nu = b_\nu^{-1} P^* \psi_\nu, \quad \nu \in N
\]

The SVD of \( P \) is therefore given abstractly by the equations \( P \phi_\nu = b_\nu \psi_\nu, \nu \in N \).

\( P^* \) induces orthogonality relations for \( \{\phi_\nu\} \) on \( B \) from those satisfied by \( \{\psi_\nu\} \) on \( D \); as shown by Davison (1983) this leads to an explicit identification of \( \{\phi_\nu\} \):

\[
\phi_{\ell m}^\alpha(x) = d_{\ell m}^\alpha e^{i\ell \theta} r^{\ell} P_{m-\ell, m+\ell}^{(\alpha-1,|\ell|)}(2r^2 - 1), \quad x = re^{i\theta}
\]

where \( P_{k,\ell}^{(\alpha,\beta)} \) denotes the \( k \)th degree Jacobi polynomial (as defined in GR, p. 1035 or Magnus et al. §5.2), and \((d_{\ell m}^\alpha)^2 = [(m + \alpha)(\alpha + k)!/[\alpha \alpha \alpha] \cdot j!]\). For \( \alpha = 1 \), we recover the Zernike polynomials \( \phi_{\ell m}(x) = \sqrt{m + 1} e^{i\ell \theta} Z_{m,\ell}^\alpha(r) \). This yields the explicit SVD for the case \( \alpha = 1 \) described in Section 5.

[To obtain (8) and \( d_{\ell m}^\alpha \) from Davison (1983, Theorem 3.2), note from (a) and (d) that

\[
\phi_{\ell m}(x) = e^{i\ell \theta} \frac{c_m^\alpha}{b_{\ell m}} \frac{1}{2\pi} \int_0^{2\pi} e^{i\beta} C_m^\alpha(r \cos \beta) d\beta, \quad x = (r \cos \theta, r \sin \theta)
\]

\[
= e^{i\ell \theta} \frac{c_m^\alpha}{b_{\ell m}} \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} e^{i\beta} C_m^\alpha(r \cos \beta) d\beta, \quad (m - \ell = 2k \text{ is even})
\]

\[
= \frac{c_m^\alpha}{b_{\ell m}} \frac{1}{\pi} (w_2 \Phi_{m,\ell})(x)
\]

in Davison’s notation (cf. p. 434.1–).]

5. Ellipsoids

An ellipsoid \( E \) in a Hilbert space \((H, \| \cdot \|_H)\) with orthogonal basis \( \{\phi_\nu\} \) is specified by a sequence of positive constants \( \{a_\nu^2\} \):

\[
E = \{f = \Sigma c_\nu \phi_\nu : \sum_\nu a_\nu^2 c_\nu^2 \|\phi_\nu\|_H^2 \leq C^2 \}.
\]

(9)
The following familiar device interprets certain such ellipsoids in terms of smoothness and integrability conditions (of Sobolev type) on $f$. Let $L$, a (differential) operator between Hilbert spaces $(H, \| \cdot \|_H, \{ \tilde{\phi}_\nu \})$ and $(K, \| \cdot \|_K, \{ \tilde{\chi}_\nu \})$, map $f = \Sigma c_\nu \tilde{\phi}_\nu$ to $L f = \Sigma c_\nu b_\nu \tilde{\chi}_\nu$. Then

$$
\| L f \|_K^2 = \Sigma c_\nu^2 \left( \frac{b_\nu \| \tilde{\chi}_\nu \|_K^2}{\| \tilde{\phi}_\nu \|_H^2} \right) \| \tilde{\phi}_\nu \|_H^2
$$

(10) and if $\{ b_\nu^2 \| \tilde{\chi}_\nu \|_K^2 \| \tilde{\phi}_\nu \|_H^{-2} \} \approx \{ a_\nu^2 \}$, the ellipsoid condition (9) amounts to a finiteness condition on $L f$. Here $\alpha_\nu \asymp \beta_\nu$ means there exists a constant $c$ such that $c^{-1} \leq \alpha_\nu / \beta_\nu \leq c$.

As will be explained below, $H = (L^2(B, \mu_1), \{ \tilde{\phi}_\nu \}), K = (L^2(B, \mu_{p+1}), \{ \tilde{\phi}^{p+1}_\nu \})$ and $L$ is a $p$th-order differential operator, say $D_x^p D_y^{p-r}$. We will look for a corresponding sequence $\{ a_\nu^2 \}$.

6. Differentiation of the singular functions

Two devices simplify computations: firstly, rescale $\tilde{\phi}_{\ell m}$:

$$
\tilde{\phi}_{\ell m}^{\alpha}(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{i\ell \phi} C_m^\alpha(x \cdot u_\phi) d\phi, \quad (\ell, m) \in N;
$$

from the definition $\phi_\nu = b_\nu^{-1} P^\alpha \psi_\nu$ we have $\tilde{\phi}_{\ell m} = c_m^{-1} b_{\ell m} \phi_{\ell m}$. Secondly, instead of $\partial / \partial x$, $\partial / \partial y$ we use

$$
D_x = \frac{1}{2} \left( \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right), \quad D_y = \frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right).
$$

Since $\{ \partial^{r} / \partial x^{s} y^{r-s}, r + s = p \}$ is a linear transform of $\{ D_x^r D_y^s, r + s = p \}$, no generality is lost.

The payoff lies in the simple differentiation formulas

$$
D_x \tilde{\phi}_{\ell m}^{\alpha} = \alpha \tilde{\phi}_{\ell-1, m-1}^{\alpha+1}, \quad D_y \tilde{\phi}_{\ell m}^{\alpha} = \alpha \tilde{\phi}_{\ell+1, m-1}^{\alpha+1}, \quad (11)
$$

which follow from the Gegenbauer identity $\frac{d}{dt} C_m^\alpha(t) = 2 \alpha C_m^{\alpha+1}(t)$. The raising of the index $\alpha \rightarrow \alpha + 1$ reflects a change of weight function and is the reason for the introduction of the measures $\mu_\alpha$ in the PET context. Note particularly the boundary cases

$$
D_x \tilde{\phi}_{-m, m}^{\alpha} = D_y \tilde{\phi}_{m, m}^{\alpha} = 0 \quad (12)
$$

(since $(D_x \phi_{-m, m}, \phi_{\ell, m-1})_{\mu_\alpha} = 0$ for $\ell \in \{ m - 1, m - 3, \ldots, -m + 1 \}$, etc.).

We switch between index systems $(\ell, m) \in N$ and $(j, k) \in Z_+^2$ as is convenient. (Recall $\ell = j-k$, $m = j + k$). With the convention that $\tilde{\phi} = 0$ if $j$ or $k < 0$, (11) and (12) become

$$
D_x \tilde{\phi}_{j k}^{\alpha} = \alpha \tilde{\phi}_{j-1, k}^{\alpha+1}, \quad D_y \tilde{\phi}_{j k}^{\alpha} = \alpha \tilde{\phi}_{j+1, k}^{\alpha+1}.
$$

If $f = \sum_{j,k \geq 0} c_{j k}^{(\alpha)} \tilde{\phi}_{j k}^{\alpha}$, we conclude

$$
D_x^r D_y^s f = \sum_{j \geq r \atop k \geq s} c_{j k}^{(\alpha)} (\alpha)_{r+s} \tilde{\phi}_{j-r, k-s}^{\alpha+r+s}, \quad (13)
$$

where $\alpha_{r+s} = (\alpha)^{(r+s)}$ and $c_{j k}^{(\alpha)} = c_{j k}^{\alpha}$. 


where, as before, \((\alpha)_p = \alpha(\alpha + 1) \cdots (\alpha + p - 1) = \Gamma(\alpha + p)/\Gamma(\alpha)\).

7. Smoothness and Ellipsoids

For the application in (JS, Section 5) we focus on the case \(\alpha = 1\) and \(r + s = p\). Cast in the form of (10), (13) leads to

\[
\|D^r_x D^s_x f\|_{\mu_{p+1}}^2 = \sum_{j \geq r \atop k \geq s} (c^{(1)}_{jk})^2 a^{2}_{jk} \|\tilde{\phi}_{j,k}\|_{\mu_1}^2,
\]

where

\[
a^{2}_{jk} = \frac{(c_m^{(1)})^2 \left( b_{j-r,k-s}^{p+1} \right)^2}{(p!)^2} = \frac{(k-s+p)!}{(j-r)! (k-s)!}.
\]

After noting that when \(0 \leq r \leq p\) and \(j \geq r, (j-r+p)!/(j-r)! \in (j+1)^p[(p+1)^{-p}, p^p]\), we finally obtain the bounds

\[
a^{2}_{jk} \in (j+1)^p(k+1)^p (p+1)^{-2p}, p^{2p} \quad j \geq r, k \geq s.
\]

Return to the PET application and expand \(f \in L^2(B, \mu_1)\) in the orthonormal basis \(\phi_\nu (= \phi_\nu^1)\) as \(\sum f_\nu \phi_\nu\), so that \(f_{jk} = c^{(1)}_{jk} \|\tilde{\phi}_{j,k}\|_{\mu_1}\). Hence if \(r + s = p\),

\[
\|D^r_x D^s_x f\|_{\mu_{p+1}}^2 \leq \sum_{j \geq r \atop k \geq s} (j+1)^p(k+1)^p|f_{jk}|^2.
\]

To summarize, we have arrived at the following:

**Theorem** The function \(f\) has \(p\) weak derivatives in \(L^2(B, \mu_{p+1})\) iff \(\|D^r_x D^s_x f\|_{\mu_{p+1}}^2 < \infty\) for all \(r, s \geq 0, r + s = p\), and this occurs if and only if

\[
\sum_{j+k \geq p} (j+1)^p(k+1)^p|f_{jk}|^2 < \infty. \quad (14)
\]

If in addition we require \(f\) to be square integrable \((f \in L^2(B, \mu))\), then the range of summation in (14) should be extended to \(\{(j, k) : j \geq 0, k \geq 0\}\).

**Remark:** For the connection between \(P^\alpha\) and the Radon transform, see the commutative diagrams in Davison (1981). In this treatment, the form of the adjoint \(P^*\) does not depend on \(\alpha\), in that of Davison, it is the forward mapping (the Radon transform) that is functionally independent of \(\alpha\).
References


