13

Simultaneously Band and Space Limited Functions in Two Dimensions, and Receptive Fields of Visual Neurons

Jonathan D. Victor
Bruce W. Knight

To Larry Sirovich, on the occasion of his 70th birthday.

ABSTRACT Functions of a single variable that are simultaneously band- and space-limited are useful for spectral estimation, and have also been proposed as reasonable models for the sensitivity profiles of receptive fields of neurons in primary visual cortex. Here we consider the two-dimensional extension of these ideas. Functions that are simultaneously space- and band-limited in circular regions form a natural set of families, parameterized by the "hardness" of the space- and band- limits. For a Gaussian ("soft") limit, these functions are the two-dimensional Hermite functions, with a modified Gaussian envelope. For abrupt space and spatial frequency limits, these functions are the two-dimensional analogue of the Slepian (prolate spheroidal) functions (Slepian and Pollack [1961]; Slepian [1964]). Between these limiting cases, these families of functions may be regarded as points along a 1-parameter continuum. These families and their associated operators have certain algebraic properties in common. The Hermite functions play a central role, for two reasons. They are good asymptotic approximations of the functions in the other families. Moreover, they can be decomposed both in polar coordinates and in Cartesian coordinates. This joint decomposition provides a way to construct profiles with circular symmetries from superposition of one-dimensional profiles. This result is approximately universal: it holds exactly in the "soft" (Gaussian) limit and in good approximation across the one-parameter continuum to the "hard" (Slepian) limit. These properties lead us to speculate that such two-dimensional profiles will play an important role in the understanding of visual processing in cortical areas beyond primary visual cortex. Comparison with published experimental results lends support to this conjecture.

Contents

1 Introduction

The understanding of the structure of neural computations, and how they are implemented in neural hardware, are major goals of neuroscience. The visual system often is used as a model for this purpose. For visual neurons, characterization of the spatial weighting of their various inputs is an important concrete step in this direction. For an idealized linear neuron, this spatial weighting—the sensitivity profile of the receptive field—fully describes the spatial integration performed by that neuron. Real neurons, including those of the retina and primary visual cortex, exhibit nonlinear combination of their inputs, but their receptive field profiles nevertheless provide a useful qualitative description of these neurons' response properties. For example, the circularly symmetric center/surround antagonism that characterizes typical retinal output neurons suggests a filtering process that removes overall luminance and other long-range correlations in the retinal image. The strongly oriented receptive field profiles encountered in primary visual cortex suggest extraction of one-dimensional features, oriented in corresponding directions.

The nervous system's design must represent a balance among multiple, often conflicting, demands. Efficiency (that is, representation of the spatiotemporal visual input with as few cells as possible, and with a firing rate that is, on average, as low as possible) may appear to be a main criterion for fitness. However, efficient schemes have hidden costs. These include the metabolic and morphological requirements of creating or decoding "efficient" representations (Laughlin, de Ruyter, and Anderson [1998]), the lack

of robustness of "efficient" representations in the face of damage to the network, the length and complexity of connections that may be required to implement an efficient scheme, and the burden of specifying these connections in genetic material. Nevertheless, views that consider a biologically motivated notion of efficiency, along with general aspects of the statistics of natural visual images as factors in shaping the nervous system, can provide a successful account of receptive field properties, both in the retinal output (Atick and Redlich [1990]) and in primary visual cortex (Field [1994]). Regardless of whether efficiency plays a dominant role in shaping receptive fields in primary visual cortex, the empirical observation remains that to a good approximation, many receptive field profiles are well-described by a Gabor function (Daugman [1985]; Jones and Palmer [1987]; Marcelja [1980]), namely, a Gaussian multiplied by a sinusoid. Because the period of the sinusoid and the width of the Gaussian envelope of neurons encountered in primary visual cortex are comparable, these Gabor functions also can be well approximated by a Gaussian multiplied by a low-order oscillatory polynomial, such as a Hermite polynomial. Consequently, it may be that simultaneous confinement in space and in spatial frequency suffices to account for the shape of receptive field profiles in primary visual cortex, and that orientation plays no special role. Without recourse to detailed analyses of coding strategy and image statistics, one can argue that such receptive fields are good choices (Marcelja [1980]) to minimize wiring length and connectivity (by their confinement in space) and to analyze textures, features, and images at a particular spatial scale (by their confinement in spatial frequency).

It is not clear how to pursue the above line of inquiry beyond primary visual (striate) cortex. In extrastriate visual areas, neuronal properties become progressively less stereotyped, receptive fields are less well localized, and characterization methods based on standard systems-analysis procedures appear to be progressively less useful. We do not propose to solve this problem here. However, it is striking to note that within primary visual cortex, receptive field profiles appear suited for processing primarily along a single spatial dimension. On the other hand, many aspects of visual processing are essentially two-dimensional. These include extraction of low-level features such as T-junctions (Rubin [2001]), curvature (Wolfe, Yee, and Friedman-Hill [1992]), texture (Victor and Brodie [1978]) and shape (Wilkinson, Wilson, and Habak [1998]; Wilson and Wilkinson [1998]), as well as higher-level processes such as letter and face identification. Moreover, recordings from individual neurons beyond striate cortex reveal evidence of fundamentally two-dimensional processing (Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996]; Tanaka, Saito, Fukada, and Moriya [1991]). Gallant's work (Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996]) is particularly provocative in that it suggests that some of these neurons are tuned to various kinds of circular symmetry.

These considerations motivate exploration of the consequences of simultaneous space- and band-limitation in two dimensions. As elaborated in the Discussion, our notion of "confinement" is distinct from that of Daugman [1985], whose analysis, based on a quantum-mechanical notion of "uncertainty," led to the Gabor functions as playing an optimal role. Our notion of confinement has several parametric variations, and the family of functions that optimally achieve simultaneous confinement in space and spatial frequency depends on them. However, these families of functions also have algebraic properties and qualitative behavior that are independent of these details. Each family includes functions that are good models for receptive field profiles in primary visual cortex (that is, they resemble Gabor functions), but also functions that are intrinsically two-dimensional. One such family of functions is the Hermite functions. These functions serve as good asymptotic approximations for the other families. Their properties suggest how circularly symmetric receptive fields can be built out of simple combinations of the receptive fields encountered in primary visual cortex, V1.

2 Results

2.1 Definitions

Following the general notation of Slepian and Pollack [1961], we consider space-limiting operators D and band-limiting operators B. Both are linear operators on functions f on the plane. We define a general space-limiting operator D with shape parameter a and scale parameter d by

$$D_{a,d} f(x,y) = \exp\left[-\left(\frac{|x|^2+|y|^2}{d^2}\right)^{\frac{1}{2}a}\right] f(x,y)\,. \tag{2.1}$$

The corresponding band-limiting operators B are most readily defined in the frequency domain, by

$$B_{a,b}\,\tilde{f}(\omega_x,\omega_y) = \exp\left[-\left(\frac{|\omega_x|^2+|\omega_y|^2}{b^2}\right)^{\frac{1}{2}a}\right] \tilde{f}(\omega_x,\omega_y)\,, \tag{2.2}$$

where

$$\tilde{f}(\omega_x,\omega_y) = \frac{1}{2\pi}\iint f(x,y)\,\exp\left[-i(\omega_x x+\omega_y y)\right]dx\,dy\,, \tag{2.3}$$

whence

$$f(x,y) = \frac{1}{2\pi}\iint \tilde{f}(\omega_x,\omega_y)\exp\left[i(\omega_x x+\omega_y y)\right]d\omega_x\,d\omega_y\,. \tag{2.4}$$

Note that in the limit $a\to\infty$ that $D_{\infty,d}$ sets f to 0 outside of $|x|^2+|y|^2 \le d^2$, and leaves f unchanged within the disk of radius d, and analogously for $B_{\infty,b}$. In equations (2.3) and (2.4), we use a symmetric form for the Fourier transform; this choice will prove convenient for a later analysis (see equation (2.108)).

2.2 Algebraic Properties

It follows immediately from the definition (equation 2.1) that $D_{a,t}D_{a,u} = D_{a,\nu}$, where $\frac{1}{t^a} + \frac{1}{u^a} = \frac{1}{\nu^a}$, and that $D_{\infty,d}$ is idempotent. This composition rule guarantees the existence of functional square roots,

$$(D_{a,t})^{\frac{1}{2}} = D_{a,\nu}, \text{where}\, \nu = 2^{\frac{1}{a}} t\,, \tag{2.5}$$

which will be important below. Corresponding relationships hold for B.

From here on, we will suppress subscripts whenever this does not lead to ambiguities. D and B are evidently self-adjoint operators. Since they do not commute, BD and DB are not self-adjoint, but combinations such as $D^m B^k D^m$ and $B^m D^k B^m$ are self-adjoint. Moreover, if ψ is an eigenvector of $D^u B D^\nu$ with eigenvalue λ, then $D^h\psi$ is an eigenvector of $D^{u+h}BD^{\nu-h}$, also with eigenvalue λ. This is because

$$D^{u+h}BD^{\nu-h}(D^h\psi) = D^h(D^u B D^\nu \psi) = D^h(\lambda\psi) = \lambda(D^h\psi). \tag{2.6}$$

In particular, the operators $D^u B D^v$ are isospectral for changes in u and ν which leave $u+\nu$ constant. These relations take a particularly simple form for D_∞ and B_∞, which are idempotent.

It now follows that BD and DB have only real eigenvalues (since they are isospectral with the self-adjoint operators $B^{\frac{1}{2}}DB^{\frac{1}{2}}$ and $D^{\frac{1}{2}}BD^{\frac{1}{2}}$). Moreover, their eigenvalues λ are necessarily between 0 and 1. $|\lambda| \leq 1$ follows from the observation that the operators D and B can only diminish the magnitude of a function, and $\lambda \geq 0$ follows from the inner-product calculation that for any function f,

$$\left(D^{\frac{1}{2}}BD^{\frac{1}{2}}f, f\right) = \left(B^{\frac{1}{2}}D^{\frac{1}{2}}f\,, B^{\frac{1}{2}}D^{\frac{1}{2}}f\right) \geq 0\,. \tag{2.7}$$

Consider now the problem of finding the function f that is most nearly preserved by successive application of space- and band-limits via D and B, respectively. One way of formulating this problem is as follows. Since application of D and B can only reduce $|f|^2 \equiv (f,f)$, we can formulate this problem as finding the function f that maximizes $|BDf|^2$ subject to $|f|^2 = 1$. We calculate

$$\max_{|f|^2=1} |BDf|^2 = \max_{(f,f)=1} (f, DB^2Df)\,. \tag{2.8}$$

The Lagrange multiplier method, applied to the right hand side of equation (2.8), indicates that $|BDf|^2$ is extremized when f satisfies the eigenvalue equation

$$DB^2Df = \lambda f\,. \tag{2.9}$$

By the remarks at the beginning of this section, the eigenfunctions of DB^2D can be obtained from the eigenfunctions of B^2D^2 by application of

D. The operator B^2D^2 is of the form BD, but with choices for the width parameters b and d reduced by a factor of $2^{1/a}$ (equation 2.5). Note that in the abrupt ($a = \infty$) case, the situation is particularly simple (Slepian and Pollack [1961]): this change in scale is trivial, and eigenfunctions of BD, operated on by D, are in turn eigenfunctions of DB.

Other than this change of scale and the application of D, the eigenfunction of BD with largest eigenvalue is the function f that is altered the least by successive application of space- and band-limits. Similarly, eigenfunctions of successively lower eigenvalue correspond to functions that, within subspaces orthogonal to eigenfunctions of higher eigenvalue, have this property.

The notion of a function that is simultaneously both space- and band-limited can be expressed in other ways, such as a function that maximizes $(DBDf, f)$ subject to $|f|^2 = 1$. By the above remarks, the solutions to the corresponding eigenvalue problems (equation 2.9) are readily obtained from the eigenfunctions of BD, and this relationship is particularly simple in the abrupt case. (Note that the one cannot merely ask for functions that maximize $(BDf, f) = 1$ subject to $|f|^2 = 1$. Since BD is not self-adjoint, this maximization does not lead to an eigenvalue problem for BD as did the extremum problem of equation (2.8).)

Thus, we direct our attention to finding the eigenfunctions of BD.

2.3 The Gaussian Case

We focus on the Gaussian ($a = 2$) case, for several reasons. First, the eigenfunctions of BD (and hence, DB) have a simple closed form. Second, the operators D and B generically have rotational symmetry, but in this case also separate in Cartesian coordinates. This leads to relationships between the eigenfunctions of the separated operators and eigenfunctions with rotational symmetry. Finally, the Gaussian case provides a reasonable approximate description of the general case, as anticipated from the slowly-varying linear operator theory of Sirovich and Knight (Knight and Sirovich [1982, 1986]; Sirovich and Knight [1981, 1982, 1985]).

We write $D_{gau,d} \equiv D_{2,d\sqrt{2}}$, an operator whose envelope is a product of one-dimensional Gaussians, each of standard deviation d. We use the analogous convention for $B_{gau,b} \equiv B_{2,b\sqrt{2}}$. Now it is convenient to define

$$D_{x,d}f(x,y) = \exp\left[-\frac{x^2}{2d^2}\right] f(x,y)\,, \tag{2.10}$$

which space-limits only along the x-coordinate, and to analogously define $D_{y,d}, B_{x,b}$ and $B_{y,b}$. $D_{gau,d} = D_{x,d}D_{y,d}, B_{gau,d} = B_{x,b}B_{y,b}$, and the operators also have the commutation relations $D_{x,d}B_{y,b} = B_{y,b}D_{x,d}$ and $D_{y,d}B_{x,b} = B_{x,b}D_{y,d}$. Thus, there will be eigenfunctions $\psi(x,y)$ of

$B_{gau,b}D_{gau,d}$ with eigenvalue λ which have the form

$$\psi(x,y) = \psi_x(x)\psi_y(y)\,, \tag{2.11}$$

where the factors $\psi_x(x)$ and $\psi_y(y)$ are eigenfunctions of the one-dimensional operators $B_{x,b}D_{x,d}$ and $B_{y,b}D_{y,d}$ with eigenvalues λ_x and λ_y, and

$$\lambda = \lambda_x\lambda_y\,. \tag{2.12}$$

2.4 One Dimension

We now consider the one-dimensional operators $D_{gau,d} \equiv D_{2,d\sqrt{2}}$, $B_{gau,b} \equiv B_{2,b\sqrt{2}}$, and suppress the subscripts. As the Fourier transform of a Gaussian is another Gaussian, we have explicitly,

$$\big(BDf\big)(x) = \frac{b}{\sqrt{2\pi}} \int f(u)\, \exp\Big(-\frac{u^2}{2d^2}\Big)\, \exp\left(-\tfrac{1}{2}b^2(u-x)^2\right) du\,. \tag{2.13}$$

Following Knight and Sirovich [1982, 1986], we seek solutions of

$$BDf_n = \lambda_n f_n \tag{2.14}$$

in the form of a Hermite polynomial h_n scaled by a factor k, multiplied by a Gaussian of standard deviation α,

$$f_n(x) = h_n(kx)\, \exp\left(-\frac{x^2}{2\alpha^2}\right), \tag{2.15}$$

and anticipate eigenvalues of the form

$$\lambda_n = \eta^{n+\frac{1}{2}}. \tag{2.16}$$

It is convenient to use Hermite polynomials that are orthogonal with respect to a Gaussian of unit standard deviation, and whose highest coefficient is unity. This convention, more convenient for what follows, is different from the standard one (Abramowitz and Stegun [1964] equation 22.2.14) for the Hermite polynomials H_n; the relationship between these conventions is

$$h_n(x) = 2^{-n/2} H_n\left(\frac{x}{\sqrt{2}}\right). \tag{2.17}$$

Under our convention, the Hermite polynomials have the generating function

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}\, h_n(x) = \exp\left(xz - \tfrac{1}{2}z^2\right). \tag{2.18}$$

The assumed form for f_n, equation (2.15), leads to a generating function for the right hand side of the eigenvalue equation (2.14):

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}\lambda_n f_n(x) = \eta^{\frac{1}{2}} \exp\left(-\frac{x^2}{2\alpha^2} + k\eta xz - \tfrac{1}{2}\eta^2 z^2\right). \tag{2.19}$$

We can also write a generating function for the left-hand side of equation (2.14):

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}(DBf_n)(x) = \sqrt{\frac{b^2}{b^2 + \frac{1}{d^2} + \frac{1}{\alpha^2}}}\, \exp\left[-\tfrac{1}{2}(b^2x^2 + z^2)\right] \exp\left[\tfrac{1}{2}\frac{(b^2x + kz)^2}{b^2 + \frac{1}{d^2} + \frac{1}{a^2}}\right]. \tag{2.20}$$

The two generating functions (equations (2.19) and (2.20)) have the same form. Equating the coefficients of $x^2, xz,$ and z^2 in the exponents, and also equating the two overall multiplicative factors, leads to constraints for the unknown quantities $\alpha, k,$ and η of equations (2.15) and (2.16). These constraints are satisfied by

$$k = b\sqrt{\frac{1-\eta^2}{\eta}}, \tag{2.21}$$

$$\alpha^2 = \frac{1 \pm \sqrt{1 + 4b^2d^2}}{2b^2}, \quad \text{and} \tag{2.22}$$

$$\eta = \frac{b^2}{b^2 + \frac{1}{d^2} + \frac{1}{\alpha^2}}. \tag{2.23}$$

Only the positive branch of equation (2.22) corresponds to eigenfunctions that approach zero for large x. The negative branch corresponds to imaginary α and to real eigenfunctions that diverge for large x. We will focus on the positive branch.

These equations can be placed in a more dimensionless form by taking $c = bd, \kappa = \frac{k}{b},$ and $\beta = \alpha b$. With these parameters,

$$\beta^2 = \frac{1 \pm \sqrt{1 + 4c^2}}{2}, \tag{2.24}$$

$$\eta = \frac{2c^2}{1 + 2c^2 + \sqrt{1 + 4c^2}} = \left(\frac{2c}{1 + \sqrt{1 + 4c^2}}\right)^2, \quad \text{and} \tag{2.25}$$

$$\kappa^2 = \frac{\sqrt{1 + 4c^2}}{c^2}. \tag{2.26}$$

After some algebra, the eigenfunctions (2.15) can be written

$$f_n\left(\frac{\chi}{k}\right) = h_n(\chi)\, \exp\left(-\tfrac{1}{4}\chi^2\right)\, \exp\left(\frac{\chi^2}{4\kappa^2c^2}\right) = h_n(\chi)\, \exp\left(-\tfrac{1}{4}\chi^2\right)\, \exp\left(\frac{\chi^2}{4\sqrt{1 + 4c^2}}\right). \tag{2.27}$$

Equation (2.15) and the more explicit form (2.27) are thus the functions that extremize simultaneous space- and band-limited functions in the

Gaussian sense that we have defined. The combination

$$c = bd \tag{2.28}$$

is a product of the space limit and the bandwidth limit. As $c \to \infty$, the Gaussian envelope in equation (2.27) becomes progressively less prominent, and the eigenfunctions approach the unmodified Hermite functions $h_n(\chi) \exp\left(-\frac{1}{4}\chi^2\right)$.

2.5 Two Dimensions: Reorganization According to Rotational Symmetry

We now consider the two-dimensional Gaussian case. As a consequence of the Cartesian separation of equation (2.11), solutions of the two-dimensional eigenvalue problem equation (2.14) may be parameterized by integers n_x and n_y, with

$$f_{n_x,n_y}(x,y) = f_{n_x}(x)\, f_{n_y}(y)\,, \tag{2.29}$$

where the factors on the right hand side are given by equation (2.27). As a consequence of equations (2.12) and (2.16), the eigenvalue associated with f_{n_x,n_y} is

$$\lambda_{n_x,n_y} = \eta^{1+n_x+n_y}\,, \tag{2.30}$$

where η is given by equation (2.23) or (2.25).

Thus, eigenvalues are identical for eigenfunctions that share a common value of $n = n_x + n_y$. These $n+1$ eigenfunctions, namely $f_{0,n}$, $f_{1,n-1}$, $\ldots$, $f_{n-1,1}$, $f_{n,0}$, are readily reorganized into new linear combinations that exhibit polar symmetry, as one would expect from the polar symmetry of the operators B and D for the Gaussian ($a = 2$) case. To calculate these linear combinations explicitly, we combine generating functions with the umbral calculus of Rota and Taylor [1994]. The main steps are: (a) defining the umbral calculus, (b) writing a generating function for products of Hermite polynomials in Cartesian coordinates, (c) using the umbral calculus to reorganize this generating function in terms of polar coordinates, and (d) matching coefficients to arrive at the desired reorganization.

The umbral calculus is essentially an algebra of polynomials in several variables. Addition in this algebra is the usual addition. Multiplication in this algebra is a nonstandard operation that will be denoted $\otimes$. This operation is defined in terms of its action on products of Hermite polynomials (which form a basis), and then is extended to all polynomials via linearity. The linearity condition is equivalent to stating that $\otimes$ and addition obey the distributive law. For Hermite polynomials with identical formal arguments, we define

$$h_m(x) \otimes h_n(x) = h_{m+n}(x). \tag{2.31}$$

For Hermite polynomials with distinct arguments, $\otimes$ acts like ordinary multiplication:

$$h_m(x) \otimes h_n(y) = h_m(x) h_n(y) \tag{2.32}$$

We use exponential notation $p^{\otimes m} = p \otimes p \otimes \ldots \otimes p$ for iterated products of any polynomial p, and we also write $h = h_1$ so that $h^{\otimes n}(x) = h_n(x)$. For example, in this notation, the generating function (2.18) for Hermite polynomials takes the form

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} h^{\otimes n}(x) = \exp\left(xz - \frac{1}{2}z^2\right). \tag{2.33}$$

Now consider a generating function $Q(z,t)$ defined by

$$Q(z,t) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{z^k}{k!} \frac{t^l}{l!} q_{k,l}(x,y), \tag{2.34}$$

where

$$q_{k,l}(x,y) = [h(x) + ih(y)]^{\otimes k} \otimes [h(x) - ih(y)]^{\otimes l}. \tag{2.35}$$

With $k = a + r$ and $l = b + s$ and application of the binomial expansion to each term of equation (2.35) we find

$$Q(z,t) = \sum_{a,b,r,s=0}^{\infty} \frac{z^{a+r}}{a!\, r!} \frac{t^{b+s}}{b!\, s!} i^r (-i)^s \, h_{a+b}(x) \otimes h_{r+s}(y). \tag{2.36}$$

Each term is of the form of equation (2.32) at this step, so $\otimes$ becomes ordinary multiplication. Equation (2.36) now can be factored into

$$Q(z,t) = \left[\sum_{a,b=0}^{\infty} \frac{z^a t^b}{a!b!} h_{a+b}(x)\right] \cdot \left[\sum_{r,s=0}^{\infty} \frac{(iz)^r (-it)^s}{r!s!} h_{r+s}(y)\right]. \tag{2.37}$$

Application of the binomial expansion collapses each of these factors:

$$Q(z,t) = \left[\sum_{m=0}^{\infty} \frac{(z+t)^m}{m!} h_m(x)\right] \left[\sum_{m=0}^{\infty} \frac{(iz-it)^m}{m!} h_m(y)\right]. \tag{2.38}$$

It now follows from the generating function for h, equation (2.18), that

$$Q(z,t) = \exp\left[x(z+t) - \frac{(z+t)^2}{2}\right] \exp\left[y(iz-it) - \frac{(iz-it)^2}{2}\right], \tag{2.39}$$

or equivalently,

$$Q(z,t) = \exp\left[(x+iy)z + (x-iy)t - 2zt\right]. \tag{2.40}$$

With the usual polar substitutions $x = R\cos\theta$ and $y = R\sin\theta$, along with $\rho = \sqrt{zt}$ and $\sigma = \sqrt{\frac{z}{t}}$, equation (2.40) becomes

$$Q\left(\rho\sigma, \frac{\rho}{\sigma}\right) = \exp\left[R\rho\left(e^{i\theta}\sigma + \frac{1}{e^{i\theta}\sigma}\right) - 2\rho^2\right]. \tag{2.41}$$

We now form a Taylor series expansion of the right hand side:

$$\begin{aligned} Q\left(\rho\sigma, \frac{\rho}{\sigma}\right) &= \sum_{s=0}^{\infty} \frac{1}{s!}\left(R\rho\left(e^{i\theta}\sigma + \frac{1}{e^{i\theta}\sigma}\right) - 2\rho^2\right)^s \\ &= \sum_{s=0}^{\infty}\sum_{g=0}^{s} \frac{1}{g!(s-g)!}\left(R\rho\left(e^{i\theta}\sigma + \frac{1}{e^{i\theta}\sigma}\right)\right)^g \left(-2\rho^2\right)^{s-g} \\ &= \sum_{s=0}^{\infty}\sum_{g=0}^{s}\sum_{j=0}^{g} \frac{1}{(s-g)!j!(g-j)!}(R\rho)^g \left(-2\rho^2\right)^{s-g}\left(e^{i\theta}\sigma\right)^{2j-g}. \end{aligned} \tag{2.42}$$

From the middle line of equation (2.42), we see that any term that involves σ^μ must have $g \geq |\mu|$, and hence must be associated with $\rho^{2\nu+|\mu|}$ for some non-negative integer ν. We therefore collect terms that involve $\rho^{2\nu+|\mu|}\sigma^\mu$ in equation (2.42). These are the terms for which $j = \frac{1}{2}(|\mu| + g)$ and $s = \frac{1}{2}(|\mu| + g) + \nu$. Thus

$$\begin{aligned} Q\big(\rho\sigma, \frac{\rho}{\sigma}\big) = \sum_{\mu=-\infty}^{\infty}\sum_{\nu=0}^{\infty} &\rho^{2\nu+|\mu|}\sigma^\mu e^{i\mu\theta}R^{|\mu|} \\ &\times \sum_g \frac{(-2)^{\frac{1}{2}(|\mu|-g)+\nu}R^{g-|\mu|}}{\left(\frac{1}{2}(|\mu|-g)+\nu\right)!\,\left(\frac{1}{2}(|\mu|+g)\right)!\,\left(\frac{1}{2}(g-|\mu|)\right)!} \end{aligned} \tag{2.43}$$

where the inner sum is over all values of g for which the arguments of the factorials are non-negative integers. With $p = \frac{1}{2}(g - |\mu|)$, we have

$$Q\left(\rho\sigma, \frac{\rho}{\sigma}\right) = \sum_{\mu=-\infty}^{\infty}\sum_{\nu=0}^{\infty} \rho^{2\nu+|\mu|}\sigma^\mu e^{i\mu\theta}R^{|\mu|}\sum_{p=0}^{\nu}\frac{(-2)^{\nu-p}R^{2p}}{(|\mu|+p)!p!(\nu-p)!}. \tag{2.44}$$

We now convert the expression in equation (2.34) for Q to polar form in another way. The definition of $\otimes$ leads to the identity

$$[h(x) + ih(y)] \otimes [h(x) - ih(y)] = h_2(x) + h_2(y). \tag{2.45}$$

This is a crucial step: the left-hand side is a product of Hermite polynomials in Cartesian coordinates, while the right-hand side depends only on the radius (as $h_2(u) = u^2 - 1$).

Repeated application of this identity to equation (2.35) yields

$$q_{k,l}(x, y) = [h_2(x) + h_2(y)]^{\otimes \min(k,l)} \otimes [h(x) \pm ih(y)]^{\otimes|k-l|}, \tag{2.46}$$

where the sign in the final term is chosen to match the sign of $k - l$.

Now consider the substitutions $\nu = \min(k, l)$ and $\mu = k - l$. As k and l each run from 0 to ∞, ν runs from 0 to ∞, and μ independently runs from $-\infty$ to ∞ (see Figure 2.1). Moreover, $k = \nu + \frac{\mu}{2} + \left|\frac{\mu}{2}\right|$ and $l = \nu - \frac{\mu}{2} + \left|\frac{\mu}{2}\right|$, so that (k, l) pairs of constant eigenvalue (constant $k + l$) correspond to constant values of $2v + |\mu|$. (This is the reason for the reorganization of terms between equations (2.42) and (2.43).) By use of the umbral identity (2.46), the expression (2.34) for $Q(z, t)$ can be transformed to

$$Q\left(\rho\sigma, \frac{\rho}{\sigma}\right) = \sum_{\mu=-\infty}^{\infty} \sum_{\nu=0}^{\infty} \frac{\rho^{2\nu+|\mu|}}{\left(\nu + \frac{\mu}{2} + \frac{|\mu|}{2}\right)!} \frac{\sigma^{\mu}}{\left(\nu - \frac{\mu}{2} + \frac{|\mu|}{2}\right)!} \times \left[h_2(x) + h_2(y)\right]^{\otimes\nu} \otimes \left[h(x) \pm ih(y)\right]^{\otimes\mu}, \quad (2.47)$$

where the sign in the final term is chosen to match the sign of μ.

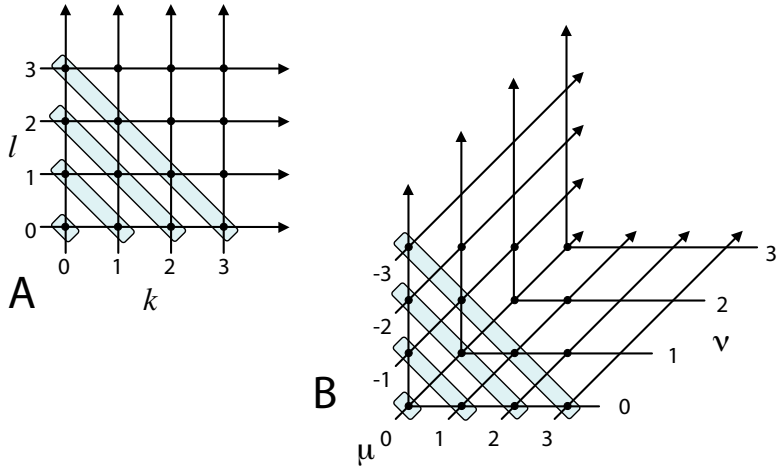

FIGURE 2.1. The change of indices from (k, l) (panel A) to (μ, ν) (panel B) via the substitutions $\nu = \min(k, l)$ and $\mu = k - l$. Coordinates that correspond to the eigenfunctions of equal eigenvalues are indicated by the gray enclosures.

We now equate the coefficient of $\rho^{2\nu+|\mu|}\sigma^{\mu}$ in equation (2.47) with the corresponding coefficient in equation (2.44). It suffices to consider $\mu \geq 0$. This yields

$$[h(x) + ih(y)]^{\otimes\mu} \otimes [h_2(x) + h_2(y)]^{\otimes\nu} = e^{i\mu\theta} R^{\mu} P_{\mu,\nu}(R^2), \quad (2.48)$$

where

$$P_{\mu,\nu}(r) = \sum_{p=0}^{\nu} (-2)^{\nu-p} \frac{(\mu+\nu)!\nu!}{(\mu+p)!p!(\nu-p)!} r^p. \quad (2.49)$$

These equations, along with equation (2.15) or (2.27) that convert the Hermite polynomials to the eigenfunctions of BD in one dimension, specify how the eigenfunctions in Cartesian separation that share a common eigenvalue can be reorganized into a polar separation. The right-hand side of equation (2.48) is separated in polar coordinates and has angular dependence $\exp(i\mu\theta)$. The radial dependence is R^μ times a polynomial $P_{\mu,\nu}$ of degree ν in R^2. Moreover, their relationship (see below) to the generalized Laguerre polynomials implies that there are ν nodes along each radius. The left-hand side of equation (2.48) is a sum of terms $h_{n_x}(x)\,h_{n_y}(y)$, as specified by the properties of $\otimes$. The indices n_x and n_y that would appear on the left- hand side in Cartesian form are all those that satisfy

$$n = n_x + n_y = 2\nu + |\mu|. \tag{2.50}$$

As an example of this reorganization, we take $\mu = 2$ and $\nu = 3$. Equation (2.48) becomes

$$[h(x) + ih(y)]^{\otimes 2} \otimes [h_2(x) + h_2(y)]^{\otimes 3} = e^{2i\theta} R^2 (R^6 - 30R^4 + 240R^2 - 480). \tag{2.51}$$

Reduction of the left-hand side via the definition of $\otimes$ leads to

$$\begin{aligned}
&[h(x) + ih(y)]^{\otimes 2} \otimes [h_2(x) + h_2(y)]^{\otimes 3} \\
&\quad = [h_2(x) + 2ih_1(x)h_1(y) - h_2(y)] \\
&\qquad \otimes [h_6(x) + 3h_4(x)h_2(y) + 3h_2(x)h_4(y) + h_6(y)] \\
&\quad = h_8(x) + 2ih_7(x)h_1(y) + 2h_6(x)h_2(y) + 6ih_5(x)h_3(y) \\
&\qquad + 6ih_3(x)h_5(y) - 2h_2(x)h_6(y) \\
&\qquad + 2ih_1(x)h_7(y) - h_8(y)\,.
\end{aligned} \tag{2.52}$$

Thus, the real part $h_8(x) + 2h_6(x)h_2(y) - 2h_2(x)h_6(y) - h_8(y)$ and the imaginary part $2h_7(x)h_1(y) + 6h_5(x)h_3(y) + 6h_3(x)h_5(y) + 2h_1(x)h_7(y)$ are the two polynomials associated with eigenfunctions of twofold axial symmetry ($\mu = 2$) and three radial nodes ($\nu = 3$). These eigenfunctions are illustrated in Figure (2.2) for $c = 4$. Figure (2.3) shows another example of this reorganization, with threefold axial symmetry ($\mu = 3$) and one radial node ($\nu = 1$), which emphasizes that the eigenfunctions in the polar separation may have symmetries manifested by none of the eigenfunctions in the Cartesian separation.

Properties of the polynomials $P_{\mu,v}$

The polynomials $P_{\mu,\nu}(R^2)$ that appear on the right-hand side of equation (2.48) are a doubly-indexed set with several interesting properties. Considered as functions on the plane, $e^{i\mu\theta} R^\mu\, P_{\mu,\nu}(R^2)$ form an orthogonal family with respect to a weight $\exp\left(-\frac{1}{2}R^2\right)$. This can be seen as follows. Two such functions that have different values of μ are orthogonal because of

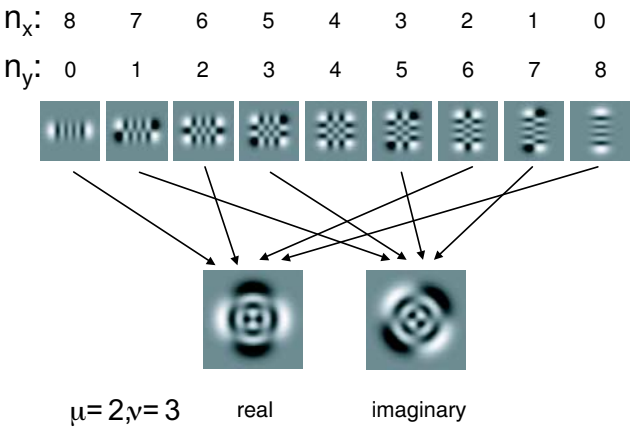

FIGURE 2.2. An example of the reorganization of eigenfunctions in the Cartesian separation into eigenfunctions in the polar separation. Top row: the nine eigenfunctions with $n_x + n_y = 8$. Bottom row: real and imaginary parts of polar separation with twofold axial symmetry ($\mu = 2$) and three radial nodes ($v = 3$), created from the Cartesian separation via equation (2.52). The space bandwidth product $c = 4$. The grayscale for each function is individually scaled.

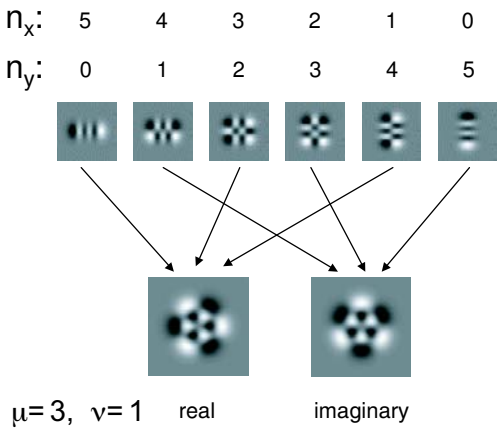

FIGURE 2.3. A second example of the reorganization of eigenfunctions in the Cartesian separation into eigenfunctions in the polar separation. Top row: the six eigenfunctions with $n_x + n_y = 5$. Bottom row: real and imaginary parts of polar separation with threefold axial symmetry ($\mu = 3$) and one radial node ($v = 1$). The space bandwidth product $c = 4$. The grayscale for each function is individually scaled.

their differing angular dependence. Two such functions that share a common value of μ but have different values of ν are orthogonal because their Cartesian decompositions are non-overlapping — since they have distinct eigenvalues (see Figure (2.1)).

Since the $e^{i\mu\theta}R^\mu P_{\mu,\nu}(R^2)$ are orthogonal on the plane with respect to $\exp\left(-\frac{1}{2}R^2\right)$, the polynomials $P_{\mu,\nu}(R^2)$ are orthogonal on the plane with respect to the weight $R^{2\mu}\exp\left(-\frac{1}{2}R^2\right)$. They thus can be considered to be generalized Hermite polynomials. Reduction by integration over circles shows that for fixed μ, these polynomials are also orthogonal with respect to a weight $R^{2\mu+1}\exp\left(-\frac{1}{2}R^2\right)$ on the half-line. With the transformation $\zeta = \frac{1}{2}R^2$, this weight becomes

$$R^{2\mu+1}\exp\left(-\tfrac{1}{2}R^2\right)dR = (2\zeta)^\mu \exp(-\zeta)d\zeta\,. \tag{2.53}$$

This demonstrates a relationship between the generalized Hermite polynomials, which have the weight on the left, and the familiar generalized Laguerre polynomials, which have the weight on the right. It extends the interrelationships expressed by equations 22.5.40 and 22.5.41 of Abramowitz and Stegun [1964].

We now find generating functions and normalization constants for these polynomials. We rewrite equation (2.41) as

$$Q\big(\rho\sigma\,,\frac{\rho}{\sigma}\big) = \exp(-2\rho^2)\,\exp\big[R\rho\big(e^{i\theta}\sigma + \frac{1}{e^{i\theta}\sigma}\big)\big] \tag{2.54}$$

and compare with the generating function for the ordinary Bessel functions $J_n(\chi)$ (equation 9.1.41 of Abramowitz and Stegun [1964]):

$$\sum_{n=-\infty}^{\infty}\tau^n J_n(\chi) = \exp\big[\tfrac{1}{2}\chi\big(\tau - \frac{1}{\tau}\big)\big]\,. \tag{2.55}$$

Taking $\tau = -ie^{i\theta}\sigma$ and $\chi = 2iR\rho$ leads to

$$Q\big(\rho\sigma\,,\tfrac{\rho}{\sigma}\big) = \exp(-2\rho^2)\sum_{n=\infty}^{\infty}\big(-ie^{i\theta}\sigma\big)^n J_n(2iR\rho)\,. \tag{2.56}$$

On the other hand, substitution of equation (2.48) into equation (2.47) gives a second expression for $Q\big(\rho\sigma,\frac{\rho}{\sigma}\big)$. Equating coefficients of $e^{i\mu\theta}$ now yields (for non-negative μ)

$$\sum_{\nu=0}^{\infty}\frac{\rho^{2\nu}}{(\nu+\mu)!\nu!}P_{\mu,\nu}(R^2) = \exp(-2\rho^2)\frac{J_\mu(2iR\rho)}{(iR\rho)^\mu}, \tag{2.57}$$

which is a generating function over ν for each series of polynomials $P_{\mu,\nu}$ with fixed ν.

To determine the normalization of the polynomials $P_{\mu,\nu}$, consider

$$\frac{1}{2\pi}\iint\sum_{0}^{\infty}\frac{z^k t^l q_{k,l}(x,y)}{k!\,l!}\frac{z'^{k'}t'^{l'}\overline{q}_{k',l'}(x,y)}{k'!\,l'!}\exp\left[-\tfrac{1}{2}(x^2+y^2)\right]dx\,dy$$
$$=\frac{1}{2\pi}\iint Q(z,t)\overline{Q}(z',t')\exp\left[-\tfrac{1}{2}(x^2+y^2)\right]dx\,dy \tag{2.58}$$

in which the sum on the left hand side is over all values of k, l, k', and l'. Via substitution of the expression (2.40) for the generating function Q and straightforward algebra, this is seen to be

$$\frac{1}{2\pi}\iint\sum_{0}^{\infty}\frac{z^k t^l q_{k,j}(x,y)}{k!\,l!}\frac{z'^{k'}t'^{l'}\overline{q}_{k',l'}(x,y)}{k'!\,l'!}\exp\left[-\tfrac{1}{2}(x^2+y^2)\right]dx\,dy$$
$$=\exp\left(2zz'+2tt'\right). \tag{2.59}$$

Consequently,

$$\frac{1}{2\pi}\iint\left|q_{k,l}(x,y)\right|^2\exp\left[-\tfrac{1}{2}(x^2+y^2)\right]dx\,dy=2^{k+l}k!\,l!\,, \tag{2.60}$$

and, as expected, cross-terms ($k\neq k'$ or $l\neq l'$) are zero. In polar form, recognizing (from equations (2.35) and (2.48)) that

$$q_{\mu+\nu,\nu}(x,y)=\exp\left(i\mu\theta\right)R^{\mu}P_{\mu,\nu}(R^2), \tag{2.61}$$

we find

$$\int_0^{\infty}\left|P_{\mu,\nu}(R^2)\right|^2R^{2\mu+1}\exp\left(-\tfrac{1}{2}R^2\right)dR=2^{\mu+2\nu}(\mu+\nu)!\,\nu!\,, \tag{2.62}$$

the normalization of the polynomials $P_{\mu,\nu}$.

2.6 The Non-Gaussian Case: One Dimension

We now return to the general band- and space-limiting operator B_aD_a, in which the profiles of the limiters are determined by the shape parameter a, as in equations (2.1) and (2.2). The eigenfunctions of the one-dimensional operator $B_\infty D_\infty$, the prolate spheroidal functions (Slepian functions, from Slepian and Pollack [1961]), closely resemble (Flammer [1957]; Xu, Haykin, and Racine [1999]) the eigenfunctions of the one-dimensional Gaussian operators B_2D_2, equation (2.27). A similar observation holds for the eigenfunctions of the corresponding two-dimensional operators (Slepian [1964]). This is rather remarkable, since the operator $B_\infty D_\infty$ limits abruptly in space and frequency, while the operator B_2D_2 applies smooth cutoffs. In this and the next sections, we provide a rationale for these similarities by drawing on the theory of Sirovich and Knight (Knight and Sirovich [1982,

1986]; Sirovich and Knight [1981, 1982, 1985]) of slowly varying linear operators. Our analysis applies not only to $B_\infty D_\infty$ but also to the operators $B_a D_a$ for intermediate exponents $a > 2$. This result can be viewed as an extension of the known asymptotic relationship between the Hermite functions and the prolate spheroidal functions (Flammer [1957]). We consider the one-dimensional case in some detail and then sketch how the arguments extend to two dimensions.

The theory of Sirovich and Knight yields asymptotic eigenvalues and eigenfunctions for integral kernels that have a slow dependence in a specific technical sense. It also delivers exact results for a broad class of kernels to which a quite general family of kernels are generically asymptotic over the principal part of their eigenspaces. It includes the WKB method for second-order differential equations as a special case. Reference Knight and Sirovich [1982] presents its application to problems related to the one here, and we now summarize the relevant part of that reference.

An integral kernel $K\{x, x'\}$ may be re-parameterized in terms of difference and mean variables

$$\nu = x - x' \quad \text{and} \quad q = \tfrac{1}{2}(x + x')\,. \tag{2.63}$$

In terms of these variables, the kernel is

$$K(\nu, q) = K\{q + \tfrac{1}{2}\nu\,, q - \tfrac{1}{2}\nu\}. \tag{2.64}$$

In the special case of a difference kernel $K\{x - x'\}$, the dependence on q in equation (2.64) is absent. The Wigner transform of $K\{x, x'\}$ is defined as

$$\begin{aligned} W(K\{x, x'\}) = \widetilde{K}(p, q) &= \int_{-\infty}^{\infty} e^{-ip\nu} K(\nu, q) d\nu \\ &= \int_{-\infty}^{\infty} e^{-ip\nu} K\left\{q + \tfrac{1}{2}\nu, q - \tfrac{1}{2}\nu\right\} d\nu\,. \end{aligned} \tag{2.65}$$

In the special case of a difference kernel $K\{x - x'\}$, equation (2.65) is simply a Fourier transform and the kernel's eigenfunctions e^{ipx} have eigenvalues $\widetilde{K}(p)$.

Note that if $K\{x, x'\}$ is symmetric in its arguments, then $\widetilde{K}(p, q)$ is real, and the implicit relation

$$\widetilde{K}(p, q) = \lambda\,, \tag{2.66}$$

for real λ, will yield a set of contour lines on the (p, q) plane. If these contour lines are *closed*, then we may pick a subset of them with specific enclosed areas:

$$\widetilde{K}(p, q) = \lambda_n\,, \quad \text{where } (p, q) \text{ encloses area } A(\lambda_n) = (2n + 1)\pi\,. \tag{2.67}$$

If $\lambda_{n+1} - \lambda_n$ is a stable small fraction of λ_n over a span of consecutive n's, then equation (2.67) gives a good estimate of the nth eigenvalue. (By

experience, taking liberties with the criterion shows the estimate is robust.) Estimated eigenfunctions also emerge from the analysis. If we solve equation (2.67) for $p_n(q)$, then we find that locally at x, the nth eigenfunction vibrates with changing x at a frequency $2\pi p_n(x)$.

In one type of circumstance, the eigenvalues of equation (2.67) are *exact*: if the contours of equation (2.66) are concentric similar ellipses and if also, in (λ, p, q) 3-space the surface (2.66) is a paraboloid. In this case, the spacing between the consecutive eigenvalues is constant. In addition, in this case the exact eigenfunctions also may be specified in terms of Hermite functions.

This exact result leads to useful asymptotics in general situations that occur commonly. If a kernel transform $\widetilde{K}(p,q)$ is expanded about an extremum (p_0, q_0), then near that extremum the paraboloidal form is generic through second-order terms in $(p - p_0, q - q_0)$. In the examples below, $\widetilde{K}(p,q)$ is symmetric under reflection of either axis. Thus it is extremized at the origin, where Taylor expansion gives

$$\widetilde{K}(p,q) \approx \widetilde{K}(0,0) + \tfrac{1}{2}\widetilde{K}_{pp}(0,0)p^2 + \tfrac{1}{2}\widetilde{K}_{qq}(0,0)q^2 \,. \tag{2.68}$$

For the right-hand expression, equation (2.67) is exact and yields exact eigenfunctions. Over the set of areas $A(\lambda_n)$ in equation (2.67) for which $\widetilde{K}(p,q)$ is well approximated by equation (2.68), the exact eigenvalues of equation (2.68) will be good estimates of those sought, and similarly for the exact eigenfunctions of equation (2.68).

These observations imply that a smooth area-preserving transformation on the (p,q) plane must map $\widetilde{K}(p,q)$ to the Wigner transform of another kernel with the same asymptotic eigenvalues given by equation (2.67). As shown in Knight and Sirovich [1982], a more restrictive *unimodular affine* transformation on (p,q), which preserves straight lines as well as areas, yields a new kernel with *exactly* the same eigenvalues. The new kernel K' is related to the original kernel by a similarity transformation T:

$$K' = TKT^{-1} \,. \tag{2.69}$$

Here, T maps the orthonormal eigenfunctions of the original kernel to those of the new, and hence preserves inner products.

A unimodular affine transformation carries an ellipse to another ellipse with equal area. Given an ellipse, a unimodular affine transformation may be constructed which carries it to a circle centered at the origin. Construction of the corresponding similarity transformation T (equation (2.69)) is also straightforward and quite simple in the axis-symmetric case (equation (2.68)). The availability of the inverse transformation, from a centered circle to an arbitrary ellipse of equal area, reduces the eigenvalue problem for a kernel whose Wigner transform yields contour lines which are concentric similar ellipses, to the eigenvalue problem for a kernel whose Wigner transform contours are origin-centered concentric circles.

Thus (still following Knight and Sirovich [1982]), we examine the eigenvalue problem for a kernel whose Wigner transform is of the form

$$\widetilde{K}(p,q) = \widetilde{K}(p^2+q^2) = \widetilde{K}(J)\,. \tag{2.70}$$

As noted above, the paraboloidal special case

$$\widetilde{K}(p,q) = a + b(p^2+q^2) = a + bJ \tag{2.71}$$

satisfies the area rule (equation (2.67)) exactly, with

$$\lambda_n = a + b(2n+1)\,. \tag{2.72}$$

For other cases of equation (2.70), the area rule is *not exact*, and we may furnish a next-order error term

$$\lambda_n = \widetilde{K}(2n+1) + \tfrac{1}{2}\widetilde{K}_{JJ}(2n+1)\,. \tag{2.73}$$

Nonetheless for kernels with Wigner transforms of the form (2.70), the eigenvalue problem may still be solved *exactly*. This may be shown (Knight and Sirovich [1982]) by relating two classical generating function formulas. The orthonormal eigenfunctions that emerge from equation (2.70) are those encountered in the limiting case ($c \to \infty$) of equation (2.27), namely the normalized Hermite functions

$$u_n(x) = \frac{1}{\sqrt[4]{\pi}}\frac{1}{\sqrt{n!}}h_n\big(\sqrt{2}x\big)\exp\big(-\tfrac{1}{2}x^2\big)\,. \tag{2.74}$$

In terms of these, the classical Mehler's formula (equation 22, section 10.3 in Erdelyi [1955]) is

$$\begin{aligned} G\{x,x'\} &= \frac{1}{\sqrt{\pi(1-z^2)}}\exp\left\{-\frac{\frac{1}{2}(z^2+1)(x^2+x'^2)-2zxx'}{1-z^2}\right\} \\ &= \frac{1}{\sqrt{\pi(1-z^2)}}\exp\left\{-\tfrac{1}{4}\big(\frac{1+z}{1-z}(x-x')^2+\frac{1-z}{1+z}(x+x')^2\big)\right\} \\ &= \sum_{n=0}^{\infty} z^n u_n(x)u_n(x')\,. \end{aligned} \tag{2.75}$$

The second line has been arranged in a form easier to compare to what's above and particularly with equation (2.63). Clearly $G\{x,x'\}$ is an integral kernel whose nth eigenfunction is $u_n(x)$ and nth eigenvalue is z^n, whence

$$Gu_n = z^n u_n\,. \tag{2.76}$$

Each term $u_n(x)u_n(x')$ on the right is a 1-dimensional projection kernel. Wigner transformation of Mehler's formula (2.75) (a straightforward "complete the squares" integral) yields

$$\widetilde{G}(p^2+q^2,z) = \frac{2}{1+z}\exp\left\{-\frac{1-z}{1+z}(p^2+q^2)\right\} = \sum_{n=0}^{\infty} z^n W\big(u_n(x)\,u_n(x')\big)\,. \tag{2.77}$$

We can expand the left-hand expression in powers of z, which shows that each projection kernel has a Wigner transform which is constant on concentric circular contour lines. The "concentric circular contours" property clearly is inherited by any weighted sum of functions of (p,q) which individually have that property. Thus a kernel of the form

$$K\{x,x'\} = \sum_{n=0}^{\infty} \lambda_n u_n(x)\,u_n(x') \tag{2.78}$$

will have a circular-contour Wigner transform as in equation (2.70). Is the converse true? Can any kernel which satisfies equation (2.70) be expressed in the form (2.78) (which solves the eigenvalue problem)? Compare equation (2.77) with the generating function for the orthonormal Laguerre functions $\mathtt{L}_n(x)$ (derived from the generating function for the standard Laguerre polynomials L_n from equation 22.9.15 of Abramowitz and Stegun [1964] with $\mathtt{L}_n(x) = (-1)^n e^{-\frac{1}{2}x} L_n(x)$):

$$\frac{1}{1+z}\exp\left\{-\tfrac{1}{2}\big(\frac{1-z}{1+z}\big)J\right\} = \sum_{n=0}^{\infty} z^n \mathtt{L}_n(J)\,. \tag{2.79}$$

We see that

$$W\big(u_n(x)\,u_n(x')\big) = 2\mathtt{L}_n(2J) \tag{2.80}$$

and these functions are a complete orthonormal set. Thus a projection integral applied to equation (2.70) evaluates the eigenvalue:

$$\lambda_n = \int_0^{\infty} \widetilde{K}(J)\cdot 2\mathtt{L}_n(2J)\,dJ\,. \tag{2.81}$$

Our converse holds because of the completeness of the Laguerre functions.

The Wigner transform of equation (2.78) is

$$\widetilde{K}(p^2+q^2) = \sum_{n=0}^{\infty} \lambda_n \cdot 2\mathtt{L}_n\big(2(p^2+q^2)\big)\,. \tag{2.82}$$

Each of the Laguerre functions here has a peaking form which gives a dominant contribution to the sum when p^2+q^2 is near n, in qualitative agreement with the area rule (2.67). Below we will encounter kernels which are associated with the non-Gaussian space- and band-limited kernels and which share their eigenfunctions. The Wigner transforms of these kernels

show a central regime of near-circular contours with rapid radial variation, and this feature will confirm their asymptotic agreement with the form of equation (2.70).

We will first apply the methodology of Sirovich and Knight to the space- and band-limited kernels themselves, which yields some insight but inconclusive results. We will then work with the more definitive associated kernels.

A first step in applying this theory is to focus on the self-adjoint operator $D^{\frac{1}{2}}BD^{\frac{1}{2}}$. We can write

$$D^{\frac{1}{2}}BD^{\frac{1}{2}}f(x) = \int K\{x,x'\}f(x')\,dx', \tag{2.83}$$

where

$$K\{x,x'\} = [D(x)]^{\frac{1}{2}}[D(x')]^{\frac{1}{2}}B(x-x')\,. \tag{2.84}$$

Here, $D(x)$ is the spatial profile which corresponds to the space-limiting operator D

$$D(x) = e^{-(|x|/d)^a}\,, \tag{2.85}$$

the one-dimensional analog of equation (2.1), and $B(x)$ is the Fourier transform of the analogous frequency-limiting profile of B, namely

$$B(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[-\left(\frac{|\omega|}{b}\right)^a\right]d\omega\,, \tag{2.86}$$

with the convention following equation (2.4) for $a = \infty$. We make the substitutions $\nu = x - x'$ and $q = \frac{1}{2}(x + x')$, with the intent of considering K as varying slowly with q or rapidly with ν. This corresponds to the limit that the space-bandwidth product $c = bd$ is large. We next calculate the Wigner transform of K:

$$\begin{aligned}\widetilde{K}(p,q) &= \int_{-\infty}^{\infty} e^{-i\nu p}K(\nu,q)\,du \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left[D\left(\tfrac{1}{2}\nu+q\right)\right]^{\frac{1}{2}}\left[D\left(\tfrac{1}{2}\nu-q\right)\right]^{\frac{1}{2}}B(\nu)e^{-i\nu p}\,d\nu\,.\end{aligned} \tag{2.87}$$

For $a = 2$, the Wigner transform is exactly a Gaussian,

$$\widetilde{K}(p,q) = \frac{bd}{\sqrt{1+b^2d^2}}\exp\left(-\frac{q^2}{d^2} - \frac{p^2d^2}{1+b^2d^2}\right). \tag{2.88}$$

(Here we have used $D = D_{2,d} \equiv D_{gau,d/\sqrt{2}}$ and similarly for B; the derivations from equations (2.13) to (2.27) used $D = D_{gau,d} \equiv D_{2,d\sqrt{2}}$.) The contour lines are concentric, similar ellipses around the origin. To follow the discussion above, under the unimodular transformation

$$q = \frac{d}{\sqrt[4]{1+(bd)^2}}\,\hat{q}\,, \quad p = \frac{\sqrt[4]{1+(bd)^2}}{d}\,\hat{p}\,, \tag{2.89}$$

equation (2.88) becomes

$$\widetilde{K}\left(\hat{p}^2+\hat{q}^2\right)=\frac{bd}{\sqrt{1+(bd)^2}}\exp\left(-\frac{1}{\sqrt{1+(bd)^2}}(\hat{p}^2+\hat{q}^2)\right). \tag{2.90}$$

This is just the general form of $\tilde{G}$ in (2.77) above. In that expression, if we let

$$z=\frac{\sqrt{1+(bd)^2}-1}{\sqrt{1+(bd)^2}+1}, \tag{2.91}$$

we see that

$$\widetilde{K}(\hat{p}^2+\hat{q}^2)=\frac{bd}{1+\sqrt{1+(bd)^2}}\,\widetilde{G}\left(\hat{p}^2+\hat{q}^2,\frac{\sqrt{1+(bd)^2}-1}{\sqrt{1+(bd)^2}+1}\right). \tag{2.92}$$

Consequently, by equation (2.75), the eigenvalues are

$$\lambda_n=\frac{bd}{1+\sqrt{1+(bd)^2}}\left(\frac{\sqrt{1+(bd)^2}-1}{\sqrt{1+(bd)^2}+1}\right)^n. \tag{2.93}$$

The exact eigenfunctions likewise may be found from equation (2.75) and from the inverse transformation of the first member of equation (2.89).

If the space-bandwidth product bd is chosen to be large, we see from equation (2.93) that for early n, a succession of eigenvalues will lie near unity. Equation (2.90) similarly shows that for large bd, the Wigner transform will be near unity for an extended neighborhood around the origin, which extends to $\hat{p}^2+\hat{q}^2=bd$.

Figure 2.4 shows relief maps of the Wigner transform $\widetilde{K}$ for a range of shape parameters ($a=1,2,4,\infty$) and values of the space-bandwidth product $c=bd$ ($c=\frac{1}{4}\pi,\pi,4\pi$). The top row of each part of Figure 2.5 shows a top-down view. We see that the asymptotic result found analytically above for $a=2$ of an extended region at an altitude near unity is already manifest at the modest value of $c=\pi$ and is more pronounced for $a>2$. In fact, for the Slepian case of $a=\infty$, the known eigenvalue spectrum has early values near unity and a sudden plunge to near zero at a critical n which depends on the space-bandwidth parameter c. By applying the area rule to the area of the plateau near unity in this case, we can get a good estimate of the critical n where the plunge occurs. However, the very flatness of the plateau reflects the non-generic feature of numerous almost-degenerate eigenvalues, and this confounds attempts to deduce the features of the eigenfunctions from the features of the contour lines. In the $a=\infty$ case this problem is particularly severe: the Wigner transform essentially involves the band-limited Fourier inversion integral of a function which suddenly jumps to zero. The consequent inevitable Gibbs phenomenon, which simply reflects the location of this jump, is the most prominent altitude feature on the otherwise almost flat plateau.

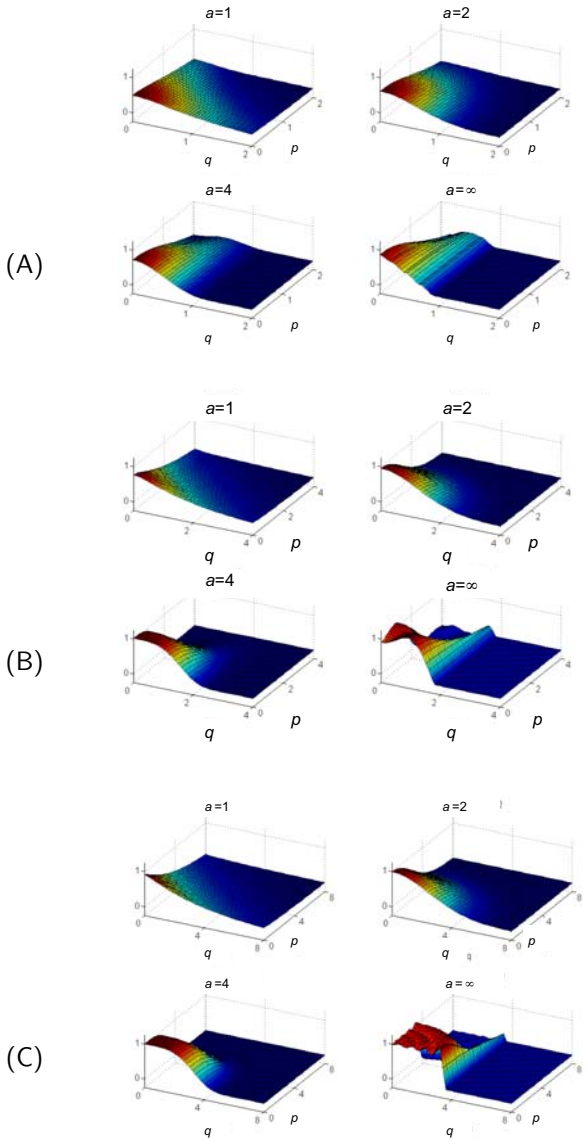

FIGURE 2.4. Wigner transforms $\widetilde{K}(p,q)$ of the operator $D^{\frac{1}{2}}BD^{\frac{1}{2}}$ for four values of the shape parameter a. The scale parameters b and d are given by $b = d = \sqrt{c}$, where $c = \frac{1}{4}\pi$ (panel A), $c = \pi$ (panel B), and $c = 4\pi$ (panel C). Only one quadrant is shown, since the transforms have even symmetry in both arguments. In each plot, the color scale runs from blue (at minimum amplitude) to deep red (at maximum amplitude).

2.7 The Non-Gaussian Case: Another Viewpoint

The above analysis is admittedly incomplete. For large values of the exponent a, the appearance of the Gibbs phenomenon complicates the asymptotic behavior of the Wigner transform away from the origin. Moreover, the presence of a nearly flat high plateau, which reflects the presence of numerous nearly degenerate eigenvalues (a non-generic feature), disrupts the straightforward application of the Sirovich–Knight methodology. These difficulties can be circumvented by an alternative approach. This approach makes explicit use of the Fourier transform relationship between the band-limiting and space-limiting operators B and D as well as the fact that taking the functional square root of either operator is equivalent to making a change in scale.

An examination of the band- and space-limiting kernel in terms of its underlying components gives some further insight into its mathematical structure. In the one-dimensional case, the natural inner product is

$$(f, g) = \int_{-\infty}^{\infty} f^*(x) g(x)\, dx\,, \tag{2.94}$$

where the asterisk indicates complex conjugation. For an operator A, its adjoint operator $A^\dagger$ is defined by

$$(A^\dagger f, g) = (f, Ag)\,. \tag{2.95}$$

The adjoint of the adjoint operator is the original operator. An operator A may be resolved as

$$A = \tfrac{1}{2}(A + A^\dagger) + \tfrac{1}{2}(A - A^\dagger) = A_R + i\, A_I\,. \tag{2.96}$$

Substitution of the definitions of A_R and A_I above (in place of A) into equation (2.95) shows that both are self-adjoint. An operator which commutes with its adjoint

$$AA^\dagger = A^\dagger A \tag{2.97}$$

is called "normal"; this is a generalization of "self-adjoint" and leads similarly to several valuable special properties. We note that equation (2.96) implies that for a normal operator, A, $A^\dagger$, A_R, and A_I all commute and have a common set of eigenfunctions. As A_R and A_I are self-adjoint, the eigenfunctions may be chosen to be orthonormal. If an eigenfunction assigns the eigenvalue $\lambda^{(R)}$ to A_R and $\lambda^{(I)}$ to A_I, then evidently the action of A upon that eigenfunction yields the eigenvalue

$$\lambda = \lambda^{(R)} + i\lambda^{(I)}\,. \tag{2.98}$$

An operator U which respects inner products

$$(Uf, Ug) = (f, g)\,. \tag{2.99}$$

is "unitary." If U and V are both unitary, clearly the concatenation UV inherits this property. We may regard the left expression of equation (2.99) as a particular case of the right expression of equation (2.95), whence

$$(U^\dagger U f, g) = (f, g)\,, \tag{2.100}$$

so

$$U^\dagger = U^{-1} \tag{2.101}$$

for a unitary operator. Since U^{-1} and U commute, a unitary operator is normal and has orthonormal eigenfunctions. In equation (2.99) we may choose both f and g to be eigenfunctions of U, and observe that the eigenvalues of U lie on the unit circle. (The above material is reviewed in many places, for example Halmos [1942]).

Now let us consider three particular operators:

(i) The parity operator P defined by

$$Pf(x) = f(-x)\,. \tag{2.102}$$

Clearly, $P^2 = 1$, whence

$$P^{-1} = P\,. \tag{2.103}$$

If (Pf, Pg) is made explicit by use of equations (2.102) and (2.94), substitution of the integration variable

$$x = -x' \tag{2.104}$$

confirms that P is unitary.

(ii) The scaling operator S_γ defined by

$$S_\gamma f(x) = \frac{1}{\sqrt{\gamma}}\, f\left(\frac{x}{\gamma}\right)\,. \tag{2.105}$$

We note

$$S_\gamma^{-1} = S_{1/\gamma}\,. \tag{2.106}$$

If $(S_\gamma f, S_\gamma g)$ is made explicit by use of equations (2.105) and (2.94), substitution of the integration variable

$$x = \gamma x' \tag{2.107}$$

shows that S_γ is unitary.

(iii) The Fourier transform operator F defined by

$$(Ff)(x) = \frac{1}{\sqrt{2\pi}} \int e^{-ixx'}\, f(x')\, dx'\,. \tag{2.108}$$

Evidently the action of the operator P on equation (2.108) exchanges $(-i)$ for $(+i)$ and yields the inverse Fourier operator. Thus,

$$PF = F^{-1} \quad \text{whence} \quad PF^2 = 1 \quad \text{and } F^2 = P\,. \tag{2.109}$$

Furthermore,

$$F^4 = 1\,, \tag{2.110}$$

so F has four eigenvalues which are the fourth-roots of unity: $i, -1, -i, 1$.

By equation (2.108), the familiar bilinear Parseval relation between a pair of functions and their Fourier transforms may be written

$$(Ff, Fg) = (f, g)\,, \tag{2.111}$$

which is an example of equation (2.99), so that F is unitary.

Direct calculation verifies that the three operators defined above have simple commutation relationships:

$$FP = PF\,, \quad PS_\gamma = S_\gamma P\,, \quad S_\gamma F = FS_{1/\gamma}\,. \tag{2.112}$$

The first two pairs simply commute, and so have a common set of eigenfunctions. The combination $F_\gamma \equiv S_\gamma F$ defined by

$$(F_\gamma f)(x) = \frac{1}{\sqrt{2\pi\gamma}} \int e^{-ixx'/\gamma} f(x')dx'\,, \tag{2.113}$$

the "scaled Fourier transform operator", will be used below to demonstrate that a set of results holds with even broader generality than is immediately apparent.

Two further properties of the Fourier operator F will prove important below. It is fairly well known (see for example Vilenkin [1968] p. 565, sec. 4, equation (1)) that the orthonormal Hermite functions $u_n(x)$ which we introduced in equation (2.74) are eigenfunctions of the Fourier operator, which order the eigenvalues we found at equation (2.110) by

$$Fu_n = (-i)^n u_n\,. \tag{2.114}$$

This is a special case of equation (2.76) for

$$z = i\,. \tag{2.115}$$

This substitution in equation (2.75) reduces G to the definition of F given in equation (2.108). The second important property of F is the expression for its Wigner transform. Direct evaluation is straightforward, or, substitution of equation (2.115) into $\widetilde{G}$ (equation 2.77) gives

$$\widetilde{F}(p^2 + q^2) = (1 + i)e^{-i(p^2+q^2)}\,. \tag{2.116}$$

Thus, $\widetilde{F}$ has a constant amplitude on the (p, q) plane, and a phase which is constant on circles and accelerates as a quadratic with increasing radius.

We next use the machinery developed above to further elucidate the structure of the operator $K\{x, x'\}$ in equation (2.83). The kernel given in equation (2.83) corresponds to the sequence of operators

$$K = D^{\frac{1}{2}} B D^{\frac{1}{2}}\,, \tag{2.117}$$

where D corresponds to simple function multiplication and B to convolution. Until further notice let us specialize by choosing the same limitation function for both space and frequency:

$$b = d = \sqrt{c} \tag{2.118}$$

for equations (2.83) and (2.84). Then, in our present notation,

$$B = F^{-1}DF \tag{2.119}$$

and so

$$K = D^{\frac{1}{2}}F^{-1}DFD^{\frac{1}{2}}\,. \tag{2.120}$$

It is convenient to adopt the notation

$$\hat{D}(x) = \big[D(x)\big]^{\frac{1}{2}} \tag{2.121}$$

and to define the convolution operator

$$\hat{B} = F^{-1}\hat{D}F\,. \tag{2.122}$$

Then B in equation (2.119) may be expressed as the iterated convolution

$$\begin{aligned} B = \big(\hat{B}\big)^2 &= \big(F^{-1}\hat{D}F\big)\big(F^{-1}\hat{D}F\big) \\ &= F^{-1}\hat{D}^2F = F^{-1}DF\,. \end{aligned} \tag{2.123}$$

Thus, equation (2.117) becomes

$$K = \hat{D}\hat{B}^2\hat{D} = \hat{D}F^{-1}\hat{D}^2F\hat{D}\,. \tag{2.124}$$

Now let us note that $\hat{D}(x)$ is an even function of x, so its multiplicative action commutes with the action of the parity operator P, defined in equation (2.102):

$$\hat{D}P = P\hat{D}\,. \tag{2.125}$$

If we insert $F^{-1} = PF$ (equation (2.109)) in equation (2.124) we find

$$\begin{aligned} K &= \hat{D}PF\hat{D}^2F\hat{D} = P\hat{D}F\hat{D}\hat{D}F\hat{D} \\ &= P\big(\hat{D}F\hat{D}\big)^2. \end{aligned} \tag{2.126}$$

Because P commutes with both $\hat{D}$ and F, from equation (2.126) we observe that the eigenfunctions of K are the same as those of its associated operator

$$Z = \hat{D}F\hat{D}. \tag{2.127}$$

As $\hat{D}$ is self-adjoint, and as F is unitary, we may now show that Z is a normal operator:

$$Z^\dagger Z = \big(\hat{D}F^{-1}\hat{D}\big)\big(\hat{D}F\hat{D}\big) = \big(\hat{D}PF\hat{D}\big)\big(\hat{D}F\hat{D}\big)$$

$$= \big(\hat{D}F\hat{D}\big)\big(\hat{D}PF\hat{D}\big) = ZZ^{\dagger} \tag{2.128}$$

and so we expect that the operator $Z = \hat{D}F\hat{D}$ will endow the eigenfunctions of K with *complex* eigenvalues. The possible separation of $\hat{D}F\hat{D}$ into a combination of two commuting self-adjoint operators (equations (2.96), (2.97)) corresponds to the possible separation of F into Fourier cosine and sine transforms,

$$F = F_R + iF_I\,. \tag{2.129}$$

The Fourier cosine component F_R is a self-adjoint operator and hence has a Wigner transform that is real. It matches F on the subspace spanned by the even-order eigenvectors and annihilates the subspace spanned by the odd-order eigenvectors. Similarly, the Fourier sine component F_I is a self-adjoint operator and has a Wigner transform that is real, and iF_I matches F on the subspace spanned by the odd-order eigenvectors.

Corresponding statements hold for the integral kernel Z. Represented as an integral kernel, $Z = \hat{D}F\hat{D}$ takes the form

$$Z\{x,x'\} = \frac{1}{\sqrt{2\pi}}\,\hat{D}(x)\,e^{-ixx'}\,\hat{D}(x')\,. \tag{2.130}$$

Noting the steps in equation (2.96), we may write

$$\begin{aligned} Z\{x,x'\} &= \frac{1}{2\pi}\big(\hat{D}(x)\cos(xx')\hat{D}(x') - i\hat{D}(x)\sin(xx')\hat{D}(x')\big) \\ &= Z_R\{x,x'\} + iZ_I\{x,x'\} \end{aligned} \tag{2.131}$$

where both Z_R and Z_I are manifestly symmetric kernels (and, as noted above, will thus have Wigner transforms which are real).

Much structural information about the operator K can be extracted from equation (2.130). From equations (2.121) and (2.85), we have that $\hat{D}(x)$ is of the form

$$\hat{D}(x) = e^{-\frac{1}{2}\left(\frac{|x|}{d}\right)^a} = e^{-\Gamma(x)}, \tag{2.132}$$

where $\Gamma(x)$ is even, zero at $x = 0$, and monotone upward to infinity. We note that as x increases, $\hat{D}(x)$ is near unity until the value of $\frac{|x|}{d}$ achieves a fair fraction of unity. If d is large, there will be a fair range of values x, x' over which $Z\{x,x'\}$ is reasonably close to $F\{x,x'\}$. As the first several eigenfunctions of F (2.74) are quite well confined to the neighborhood of the origin by their quadratic exponential factor, there is room to suspect that they might well-approximate the near-the-origin eigenfunctions of Z (which are those of K as well). We note that this was indeed the case for the "soft" Gaussian-based limiter kernel whose eigenfunctions and eigenvalues were derived exactly above. In the "hard" limit of Slepian this is likewise true: the Slepian eigenfunctions satisfy the second-order "prolate spheroidal" ordinary differential equation, which for large bandwidth

becomes asymptotic to the "parabolic cylinder" ordinary differential equation whose eigenfunctions are the Hermite functions. This is elaborated by Flammer [1957] (and has been exploited by Xu, Haykin, and Racine [1999] for the reduction of electroencephalographic data). We now pursue the conjecture for those frequency- and space-limiting kernels that lie between the "soft" and "hard" limits.

We have seen that the eigenvalue spectrum of the kernel K must lie between 1 and 0. We further noted that K commutes with the parity operator P (equations (2.112),(2.126)). Thus the eigenfunctions of K have even or odd symmetry, and assign to P the eigenvalues ± 1 respectively. In both the Slepian and Gaussian cases these eigenfunctions, unsurprisingly, alternate in parity with descending eigenvalues of K. When we factor P from K (equation (2.126)) the remaining operator Z^2 thus must have real eigenvalues which are positive or negative according to the eigenfunction's parity. Consequently the eigenvalue equation

$$Z\varphi_n = \zeta_n \varphi_n \tag{2.133}$$

must have eigenvalues which are positive or negative real for even parity, and are pure imaginary for odd parity. If we consider a sequence of space- and bandwidth-limiting operators K for which the bandwidth goes to infinity, we have

$$\hat{D} \to 1 \quad \text{and} \quad Z \to F\,. \tag{2.134}$$

This establishes that in the same limit the eigenvalues of Z go to $\{\pm 1, \pm i\}$ though it does not yet establish the choice of ± 1 on the even eigenfunctions nor $\pm i$ on the odd ones; and does not establish the eigenfunctions (2.74), because linear combinations with a common eigenvalue have not been ruled out. However, the exact solution for K in the "soft" (Gaussian) case does yield the eigenfunctions of equation (2.74) in the limit and consequently the eigenvalue sequence of equation (2.114).

The Wigner transform of the kernel Z is

$$\begin{aligned}
\widetilde{Z}(p,q) &= \int_{-\infty}^{\infty} \hat{D}\big(q+\tfrac{1}{2}\nu\big)\,\hat{D}\big(q-\tfrac{1}{2}\nu\big)\,\frac{e^{-i\left(q^2-\frac{1}{4}\nu^2+p\nu\right)}}{\sqrt{2\pi}}\,d\nu \qquad (2.135)\\
&= e^{-i(p^2+q^2)} \int_{-\infty}^{\infty} \hat{D}\big(q+\tfrac{1}{2}\nu\big)\hat{D}\big(q-\tfrac{1}{2}\nu\big)\frac{e^{i\left(\frac{1}{2}\nu-p\right)^2}}{\sqrt{2\pi}}\,d\nu\\
&= e^{-i(p^2+q^2)} \int_{-\infty}^{\infty} \hat{D}\big(q+p+\tfrac{1}{2}\nu'\big)\hat{D}\big(q-p-\tfrac{1}{2}\nu'\big)\frac{e^{i\left(\frac{1}{2}\nu'\right)^2}}{\sqrt{2\pi}}\,d\nu'\,,
\end{aligned}$$

where on the second line the square has been completed in the exponent, and on the third line $\nu = \nu' + 2p$ has shifted the integration origin; these are the two critical steps in the closed evaluation of $\widetilde{F}$ from F. In the last form of equation (2.135), the oscillations of the exponential term in the integrand accelerate as ν' advances, and eventually their signed contributions to the

integral cancel strongly. If $\hat{D}(x)$ is slow enough (from a large enough choice of the cutoff d), then for fixed (p, q), this will happen before either $\hat{D}$ factor in equation (2.135) departs enough from unity to lend a substantial (p, q)-dependence to the integral. More specifically, for large enough d, equation (2.132) may be approximated by

$$\hat{D}(x) \approx 1 - \tfrac{1}{2}\Big(\frac{|x|}{d}\Big)^a, \tag{2.136}$$

whose substitution in equation (2.135) yields

$$\widetilde{Z}(p,q) \approx \widetilde{F}(p^2+q^2) - \frac{1}{d^a\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{is^2} \left\{ |s+p+q|^a + |s+p-q|^a \right\} ds \tag{2.137}$$

(where we have let $\nu' = 2s$). As the integral on the right is a function of only p, q, and a, once any choice of these has been made, a large enough choice of the cutoff d will make $\widetilde{Z}(p, q) \to \widetilde{F}(p^2 + q^2)$. For what follows, we recall that $d = \sqrt{c}$ relates our cutoff to the space-bandwidth product.

The second and third rows of Figure 2.5 examine the limits of validity of this asymptotic approximation. Again for the modest value of $c = \pi$, and for $a \geq 2$, an amplitude plateau appears surrounding the origin and has an altitude near $\sqrt{2}$, consistent with equation (2.116). The circularity of phase contours near the origin, for $a = 4$, persists from orange through dark blue, corresponding to a half cycle. By equation (2.114), this half cycle signals dominant contributions from the first two eigenfunctions. These confirming features remain stable as we proceed to $c = 4\pi$. (Note that it is the radial scale that has been changed in Figure 2.5 to accommodate the appearance of further significant features.) However, for $a = 4$ (compare the amplitude picture with the previous) the area of the plateau has increased by a factor of 4, indicating that the number of eigenfunctions which make major contributions to this region's concordance with equation (2.116) has increased from 2 to 8. The corresponding phase picture shows that in this region both radial frequency and phase agree with equation (2.116) as well. The $a = \infty$ (Slepian) case also now shows a region of clear concurrence. Though a little shaky (when compared to $a = 4$), it extends through at least $2\frac{1}{4}$ cycles, or through the first 9 eigenfunctions. As mentioned above, as $c = bd$ increases, the early Slepian functions are *known* to approach the Hermite functions (Flammer [1957]). The figure's better concordance for an intermediate value of the shape parameter a (which supports the discussion of the asymptotics of equation (2.137)) is strong evidence that across the range of values of a, the eigenfunctions are asymptotic to the Hermite functions.

Return to the general self-adjoint operator $(D_d)^{\frac{1}{2}} B_b (D_d)^{\frac{1}{2}} = \hat{D}_d B_d \hat{D}_d$, where D_d and B_b are defined by the one-dimensional analogs of equations (2.1) and (2.2), respectively. To show that the above considerations apply even if $b \neq d$, we reconsider the scaled Fourier transform operator F_γ, as

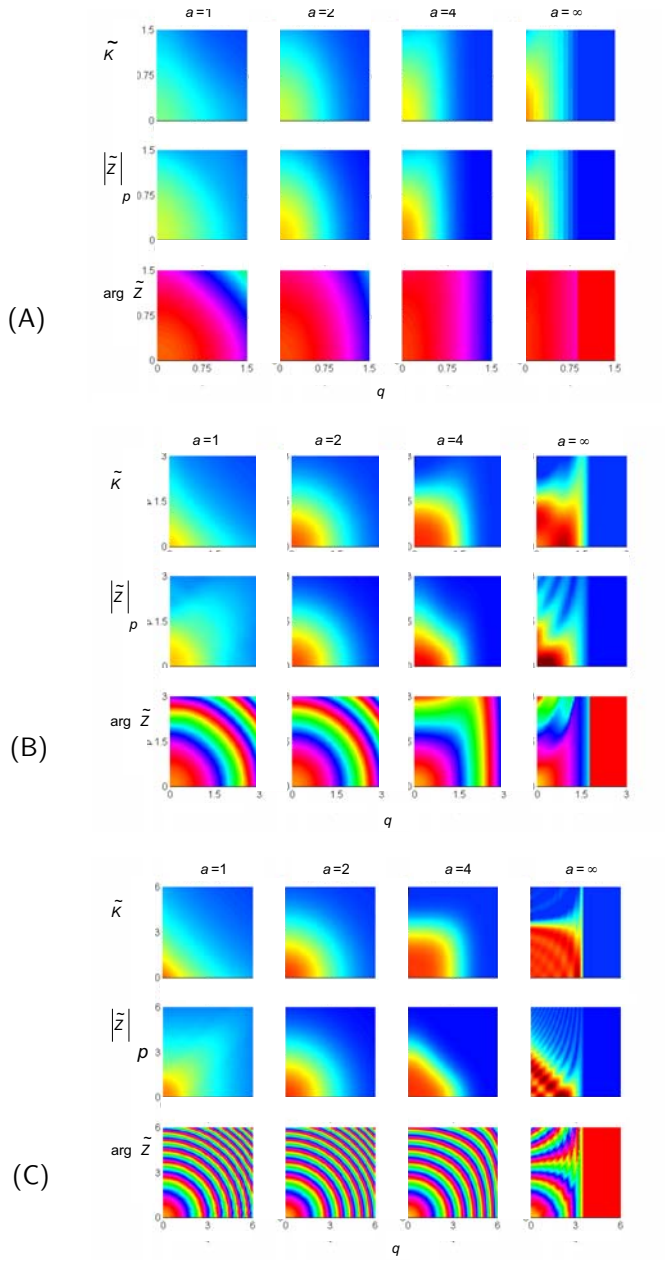

FIGURE 2.5. Comparison of the Wigner transforms $\widetilde{K}(p,q)$ and $\widetilde{Z}(p,q)$ for the values of the shape parameter a and scale parameters b and d of Figure 2.4. In each panel, first row: $\widetilde{K}(p,q)$; second row: amplitude of $\widetilde{Z}(p,q)$; third row: phase of $\widetilde{Z}(p,q)$. A common color scale for amplitude, running from blue to deep red, is used for the first two rows. Color scale for third row: red corresponds positive real, yellow to positive imaginary, green to negative real, blue to negative imaginary.

defined by equation (2.113). It follows from the definitions of B and D that

$$B_b = PF_\gamma D_{b\gamma} F_\gamma = (F_\gamma)^{-1} D_{b\gamma} F_\gamma \,, \tag{2.138}$$

generalizing equation (2.119). Hence, if we choose

$$\gamma = \frac{d}{b} \,, \tag{2.139}$$

we find

$$\hat{D}_d B_d \hat{D}_d = \hat{D}_d P F_\gamma D_d F_\gamma \hat{D}_d \,. \tag{2.140}$$

Symmetry considerations again require that any eigenfunction ψ of $\hat{D}_d B_b \hat{D}_d$ is either even- or odd-symmetric. Thus, if ψ is an eigenfunction of $\hat{D}_d B_b \hat{D}_d$

$$\hat{D}_d B_b \hat{D}_d \psi = P Z^2 \psi, \tag{2.141}$$

where we have defined

$$Z = \hat{D}_d F_\gamma \hat{D}_d \,, \tag{2.142}$$

generalizing equation (2.127).

We now proceed to find the Wigner transform $\widetilde{Z}$ of Z, where

$$Z\{x, x'\} = \frac{1}{\sqrt{2\pi\gamma}} \hat{D}_d(x) \hat{D}_d(x') e^{-ixx'/\gamma} \,. \tag{2.143}$$

The substitutions $\nu = x - x'$ and $q = \frac{1}{2}(x + x')$, followed by Fourier transformation with respect to ν, leads to an expression for the Wigner transform of Z,

$$\widetilde{Z}(p,q) = \frac{1}{\sqrt{2\pi\gamma}} \int_{-\infty}^{\infty} \hat{D}_d \left(q + \frac{1}{2}\nu \right) \hat{D}_d \left(q - \tfrac{1}{2}\nu \right) e^{-i\left(q^2 - \frac{1}{4}\nu^2\right)/\gamma} e^{-i\nu p} \, d\nu \,, \tag{2.144}$$

which is rearranged to

$$\widetilde{Z}(p,q) = \frac{1}{\sqrt{2\pi\gamma}} e^{-i(q^2/\gamma + p^2\gamma)} \int_{-\infty}^{\infty} \hat{D}_d \left(q + \tfrac{1}{2}\nu \right) \hat{D}_d \left(q - \tfrac{1}{2}\nu \right) e^{i(\nu - 2p\gamma)^2/4\gamma} \, d\nu \,, \tag{2.145}$$

generalizing equation (2.135). For $a = 2$, the integral can be performed exactly:

$$\widetilde{Z}(p,q) = \sqrt{\frac{2}{\frac{1}{bd} - i}} \exp\left(-\frac{q^2}{d^2}(1 + ibd) - \frac{p^2 d^2}{1 - ibd} \right) . \tag{2.146}$$

In the limit of $a \to \infty$, the second derivatives of the integral (2.145) at the origin behave similarly to those of $\widetilde{K}$: $\frac{\partial^2 \widetilde{Z}}{\partial p^2}$ is analytic, $\frac{\partial^2 \widetilde{Z}}{\partial p \partial q}$ is zero, and $\frac{\partial^2 \widetilde{Z}}{\partial q^2}$ is undefined. Thus, a Gaussian approximation for $\widetilde{Z}$ is no more justified than for $\widetilde{K}$. However, as in the special case of $b = d$ considered

above, the integral (2.145) can be approximated directly for the general $a > 2$, if, within the vicinity of $\nu = 2p\gamma$, the functions D_d are slowly varying. This condition translates to

$$\left|q \pm p\gamma\right| \ll 2^{1/a}\, d\,. \tag{2.147}$$

(The factor $2^{1/a}$ originates from the square root of the operator D_d.) This also takes the more symmetric form

$$\left|\frac{q}{\sqrt{\gamma}} \pm p\sqrt{\gamma}\right| \ll 2^{1/a}\sqrt{bd}\,. \tag{2.148}$$

The above condition is satisfied if both $|q| \ll d$ and $|p| \ll b$, that is, within the band- and space-limits specified by D and B, respectively. In this regime, the dominant contribution to the integral (2.145) can be obtained by replacing D with its peak value, 1. This leads to

$$\widetilde{Z}(p,q) \approx (1+i)e^{-i(q^2/\gamma + p^2\gamma)}\,, \tag{2.149}$$

generalizing the approximation of equation (2.135) by equation (2.116).

The approximate Wigner transform, equation (2.149), has contour lines that are ellipses parallel to the (q,p) coordinate axes. Moreover (similar to equation (2.89) above), these contour lines can be made circular by the symplectic transformation $q' = \frac{q}{\sqrt{\gamma}}, p' = p\sqrt{\gamma}$. Z is not Hermitian, but it is normal — that is, it commutes with its adjoint. As we have seen above, this suffices for the analysis of Knight and Sirovich [1986] to be applicable. In particular, we can conclude that within the regime specified by equation (2.148), the eigenfunctions of Z, and hence of $\hat{D}_d B_d \hat{D}_d$,are approximated by the Hermite functions.

Thus, we have generalized the asymptotic relationship of Flammer [1957] between the Hermite functions and the Slepian functions (which extremize space and band limits in the sense of $a = \infty$) to functions which extremize space and band limits for any choice of the shape parameter a in $(0, \infty)$. As the next section shows, certain features of our analysis apply even more generally, and in particular do not require specification of the shape for the space or band-limiter.

A Perturbation Analysis

For the particular case of $a = 2$ (the Gaussian case), we have an exact solution for the eigenvalues and eigenfunctions of Z. We have seen that the spectrum separates naturally into four subsets which respectively are given by $\{1, -i, -1, i\}$, each multiplied by a set of positive numbers less than unity and descending to zero. Since $K = PZ^2$ (equations (2.126) and (2.127)), this characterization of the eigenfunctions and eigenvalues of Z determines the behavior of the eigenfunctions and eigenvalues of K.

Here we use a perturbation argument to show that this behavior is generic. In contrast to the previous arguments, this aspect of the analysis is rather general, and does not depend on the precise form of the space-limiting function D or the band-limiting function B, provided that they are even-symmetric and Fourier transforms of each other.

Under these conditions, it suffices to consider the integral kernel of equation (2.130), $Z = \hat{D}F\hat{D}$, which can be recast in the more general form

$$Z\{x, x'\} = \frac{1}{\sqrt{2\pi}}\, e^{-\Gamma(x,t)} e^{-ixx'} e^{-\Gamma(x',t)}\,, \tag{2.150}$$

where t is a parameter that smoothly specifies the shape of the space-limiting function D (or $\hat{D}$), and Γ is real. In particular, we can specify that $t = 0$ corresponds to the Gaussian ("soft") case, and $t = 1$ corresponds to the Slepian ("hard") case, but for the present argument, the shapes specified by intermediate values of t need not be specified. Several choices of the parameterization by t are of interest, which includes allowing t to control a or d, the parameters that already appear in the definition (2.85) of D. Our goal is to determine how the eigenvalues ζ_n of equation (2.133) depend on t.

Since the eigenvectors φ_n may be chosen to be orthonormal, it follows from the eigenvalue equation (2.133) that

$$\zeta_n = (\varphi_n\,, Z\varphi_n)\,. \tag{2.151}$$

Consequently,

$$\frac{d\zeta_n}{dt} = -\big(\varphi_n\,, (\Gamma_t Z + Z\Gamma_t)\varphi_n\big) = -2\zeta_n(\varphi_n\,, \Gamma_t\varphi_n)\,, \tag{2.152}$$

where $\Gamma_t \equiv \frac{\partial}{\partial t}\Gamma(x,t)$. (We have used $(\varphi_n\,, \frac{\partial}{\partial t}\varphi_n) = 0$ which follows from orthonormality.) The differential equation (2.152) is readily integrated to obtain an expression for the eigenvalues ζ_n:

$$\zeta_n(t) = \zeta_n(0) e^{-2\int_0^t c_n(t')\,dt'}\,, \tag{2.153}$$

where

$$c_n(t) = \big(\varphi_n(t)\,, \Gamma_t\varphi_n(t)\big)\,. \tag{2.154}$$

Thus, as the kernel Z of equations (2.130) and (2.150) varies parametrically with t, its eigenvalues vary according to equation (2.153).

The integral in equation (2.153) is real. This means that the eigenvalues $\zeta_n(t)$ necessarily have the same phase as $\zeta_n(0)$. By choosing $\hat{D} = 1$ at $t = 0$, (so $Z = F$, the infinite space-bandwidth limit) we can thus infer from equation (2.114) that the phase of $\zeta_n(t)$ is equal to $-\frac{1}{2}n\pi$ for the general Z. Finally, we note that if the parametric dependence of γ on t is monotone increasing (corresponding, for example, to space-limiting functions $\hat{D}$ that are successively narrower), then equation (2.153) implies that as t increases, each of the eigenvalues $\zeta_n(t)$ shrinks monotonically towards zero along a cardinal axis.

2.8 The Non-Gaussian Case: Two Dimensions

The above analysis of the operator Z can be applied directly to the two-dimensional operators $D^{\frac{1}{2}}BD^{\frac{1}{2}}$. This implies a corresponding asymptotic relationship between the eigenfunctions for the general operators $D^{\frac{1}{2}}BD^{\frac{1}{2}}$ (or BD) and those for the Gaussian case.

In two dimensions, the operators B, D, and their products do not separate in Cartesian coordinates for $a \neq 2$. However, they do have exact rotational symmetry (including $a \neq 2$). This symmetry, and the separation into polar coordinates that it implies, means that one can find a complete set of eigenfunctions, all of which have an angular dependence of the form $e^{i\mu\theta}$ for some integer μ. Conversely, for each integer μ, we anticipate a discrete set of eigenfunctions $\psi_{\mu,\nu}(R,\theta) = e^{i\mu\theta}\zeta_{\mu,\nu}(R)$, where ν counts the number of zero-crossings in the radial dependence $\zeta_{\mu,\nu}(R)$. In the Gaussian case, the corresponding eigenvalue $\lambda_{\mu,\nu}$ depends only on $2\nu + |\mu|$. This degeneracy is a consequence of the dual separation into polar and Cartesian coordinates, equation (2.50).

In sum, the above analysis shows that for the general shape parameter a, the two-dimensional functions are also parameterized by a non-negative integer ν that counts the number of zero-crossings along a radius, and a second integer μ (which may be negative) that describes the angular dependence. In view of the asymptotic relationship of the general one-dimensional functions and the Hermite functions, we anticipate that the eigenvalue associated with ν and μ asymptotically depends only on $2\nu+|\mu|$.

2.9 V4 Receptive Fields

One of our motivations for studying functions that are simultaneously space- and band-limited is to gain some insight into neural processing of visual images beyond primary visual cortex (V1), and especially in V4. The extrastriate area V4 appears to be an important area for the analysis of shape Merigan [1996]. The behavior of neurons in V4 differs dramatically from those of its inputs, V1 and V2. Neurons in V1 and V2 usually can be characterized as having a single preferred orientation, and respond well to a grating of an appropriately chosen spatial frequency, especially if the grating is limited in spatial extent. Known differences between V1 and V2 neurons are relatively subtle (Levitt, Kiper, and Movshon [1994]), and consist mainly of differences in overall receptive field size, contrast sensitivity, and the prominence of responses to illusory contours (Grosof, Shapley, and Hawken [1993]; von der Heydt, Peterhans, and Baumgartner [1984]). In contrast, V4 neurons do not have a clear orientation preference, are often difficult to stimulate with gratings, and typically require complex stimuli for strong responses (Kobatake and Tanaka [1994]).

Consequently, systematic study of the spatial structure of the receptive

fields of V4 neurons has been difficult. One of the few successful examples is the provocative study by Gallant, Braun, and Van Essen [1993] and Gallant, Connor, Rakshit, Lewis, and Van Essen [1996]. V4 neurons were stimulated with patches of standard ("Cartesian") gratings, and also patches of "polar" gratings and patches of "hyperbolic" gratings (Figure 2.6). The polar gratings include target- and pinwheel-like stimuli; the hyperbolic gratings consist of alternating bands of light and dark that form rectangular hyperbolae with shared asymptotes. The main finding was that many V4 neurons were well-stimulated by particular examples of non-Cartesian stimuli. Although some neurons responded substantially better to members of one class of gratings than to the other two, most neurons had broad tunings across these classes (that is, they responded almost as well to some stimuli of the non-preferred class as to the best stimulus). Moreover, the typical neuron responded relatively poorly to most stimuli, even within its preferred class.

To determine whether this kind of behavior is expected from neurons that filter the image in a manner that is both space and band-limited, we performed a crude simulation. Receptive field profiles were constructed from the two-dimensional eigenfunctions for BD (Gaussian case), for values of the circular symmetry index $\mu \in \{0, 1, 2, 4\}$ and for a range of radial nodes $\nu \in \{0, 1, 2, 3\}$. We chose the space-bandwidth product $c = bd = 4$. The resulting profiles are shown in Figure 2.7. Raw "responses" were calculated simply by evaluating the inner product of each stimulus in Figure 2.6 with the receptive field profile. (The stimulus was considered to have a real part, as specified by Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996], and a corresponding imaginary part, obtained by replacing cosines by sines. Thus, this highly schematized model essentially posits a quadrature pair operation that removes the effect of the spatial "phase" of the Cartesian or non-Cartesian grating.) For each receptive field profile, the raw responses were then normalized by the largest quadrature pair response encountered to any of the stimuli.

A histogram of the normalized responses obtained for each receptive field profile is shown in Figure 2.8. Several features are immediately apparent. As in the data of Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996], most receptive fields responded poorly to most stimuli – as evidenced by the large peaks in most of the histograms at a response size of 0. The receptive fields corresponding to $\mu = 1$ responded almost exclusively to Cartesian gratings. This is a consequence of the symmetry of these receptive fields. Among the other profiles, many ($\mu = 4$) had optimal responses to polar gratings, and one ($\mu = 0, \nu = 1$) had an optimal response to a hyperbolic grating. The profiles corresponding to $\mu = 2$ tended to have large responses to some Cartesian and some polar gratings, while the profiles corresponding to $\mu = 4$ tended to have largest responses to polar gratings, with next-largest responses to hyperbolic gratings.

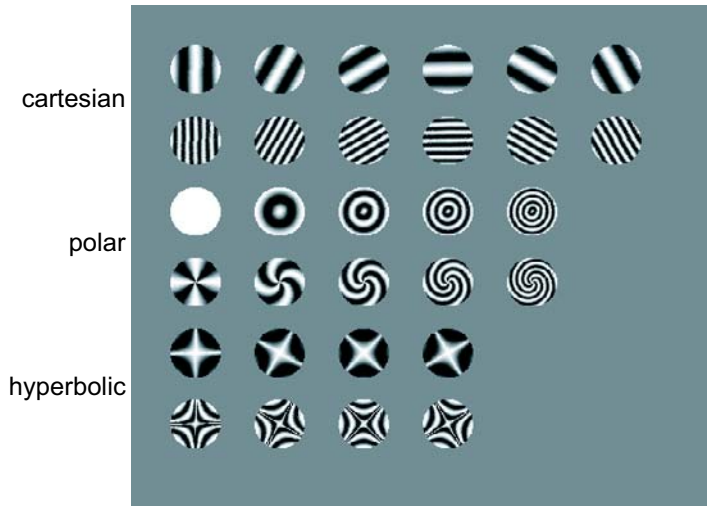

FIGURE 2.6. Some of the stimuli used by Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996]. The full stimulus set included three other spatial frequencies for each stimulus class, and also counter-clockwise variants of the "polar" stimuli — for a total of 30 Cartesian gratings, 45 polar gratings, and 20 hyperbolic gratings.

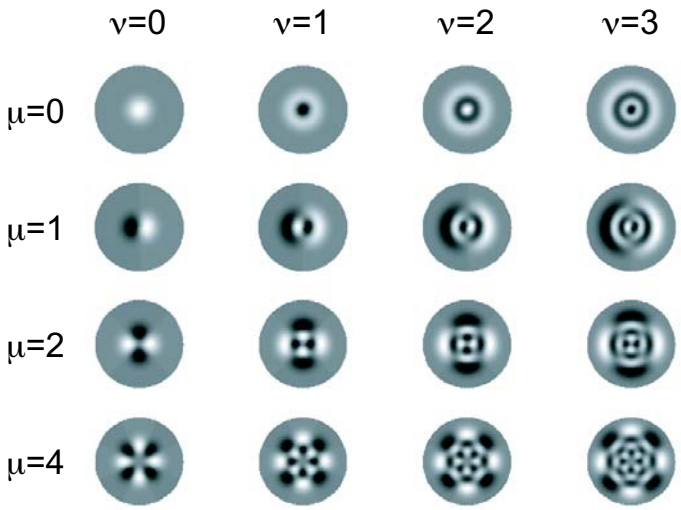

FIGURE 2.7. Receptive field sensitivity profiles (real parts) used for the simulation. All profiles had a space bandwidth product $c = 4$. The size of the disk corresponds to the region of space that covered the test stimuli, taken from Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996] (see Figure 2.6).

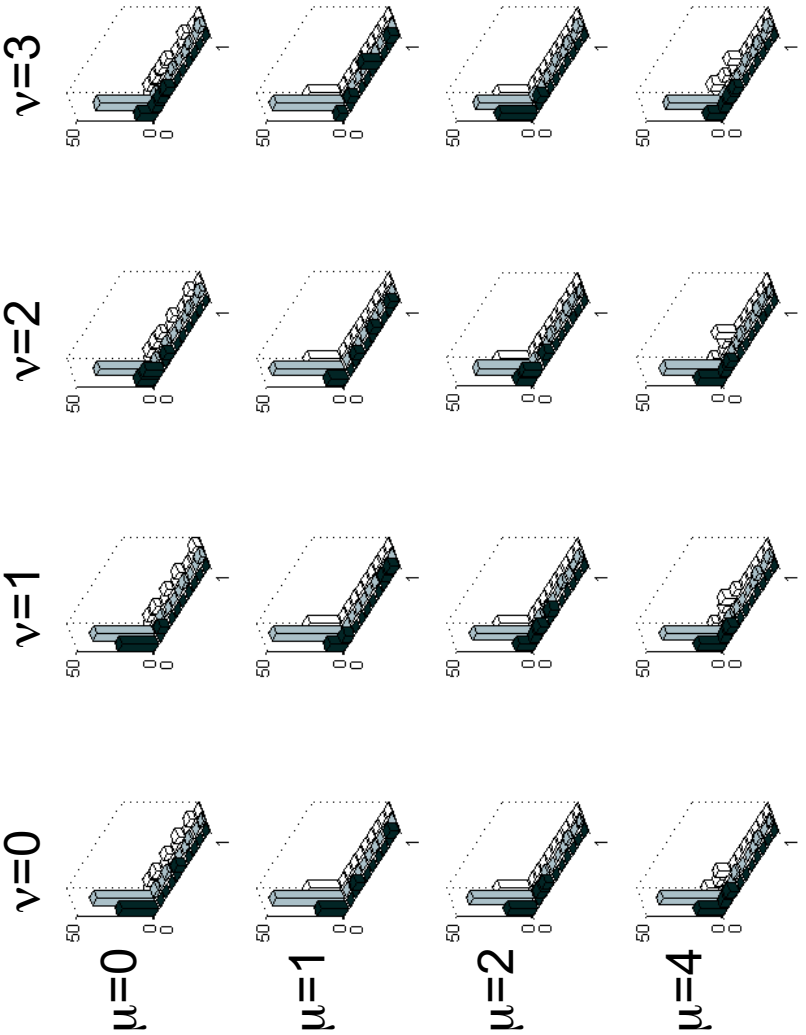

FIGURE 2.8. Summary of responses of the receptive fields of Figure 2.7 to Cartesian, polar, and hyperbolic gratings. The three histograms in each frame represent responses to the three classes of grating stimuli: Cartesian gratings (30 total), dark; polar gratings (45 total), gray; hyperbolic gratings (20 total), light. The horizontal axis represents response size, normalized by the largest quadrature pair response across all three classes. The height of each histogram bar represents the number of stimuli that elicited a response of each (normalized) size.

3 Discussion

This paper describes two-dimensional profiles that are simultaneously limited in spatial extent and in bandwidth. The main motivation is to generalize ideas that have benefited the study of visual processing in primary visual cortex, V1. As outlined above, V1 neurons, to a first approximation, are often thought of as one-dimensional spatial analyzers. They generally have a well-defined orientation preference, and typically they can be characterized by their responses to windowed gratings. However, there are indications that visual processing by individual neurons, even in V1, is not restricted to a single orientation (Purpura, Victor, and Katz [1994]). Moreover, certain aspects of early visual processing (whose physiologic locus is uncertain) make explicit use of two-dimensional information, such as extraction of T-junctions (Rubin [2001]), curvature (Wolfe, Yee, and Friedman-Hill [1992]), texture (Victor and Brodie [1978]) and shape (Wilkinson, Wilson, and Habak [1998]). Finally, neurons beyond V1 and V2 are generally not well-stimulated by standard gratings or other simple stimuli, which makes quantitative study of such neurons a challenge (Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996]; Kobatake and Tanaka [1994]; Tanaka, Saito, Fukada, and Moriya [1991]).

Our notion of simultaneous confinement in space and spatial frequency is distinct from the notion in Daugman [1985] of minimizing uncertainty in two important respects. First, (see equations (2.8) and (2.9) and surrounding material), our notion seeks profiles that are least altered by application of linear operators that truncate influences that are non-local in space (D) or spatial frequency (B). In contrast, Daugman's (1985) notion directly adopts Gabor's (1946) idea of seeking profiles that minimize joint spread in space and spatial frequency, as quantified by a Heisenberg uncertainty product. Secondly, our approach considers the profiles themselves as the subject for extremization. Though the profiles naturally group into real and imaginary pairs (lower portions of Figure 2.2 and Figure 2.3), they are intrinsically real, and their positive and negative lobes are intrinsic to their extremal properties. Daugman's approach focuses on the magnitude-squared of the profiles (Stork and Wilson [1990]); the positive and negative lobes of the derived profiles only arise because of phase factors $e^{i\omega x}$ that are irrelevant to the extremal properties.

Interestingly, applying the Gabor/Daugman approach to real-valued functions (Gabor [1946]; Stork and Wilson [1990]) identifies one-dimensional Hermite functions as playing an extremal role, but this role proves to be that of maximizing uncertainty within the space of polynomials multiplied by Gaussians, rather than minimizing it (Klein and Beutter [1992]).

The families of functions presented here have attributes that recommend them as a natural extension of windowed gratings. Since each family forms a complete orthogonal set, its members can serve as a basis for characterization of the quasilinear aspects of a neuron's response, either as in-

dividual stimuli presented in discrete trials or as members of a rapidly-presented stimulus sequence (Ringach, Sapiro, and Shapley [1997]). More importantly, these functions can serve as a common stimulus set to determine how spatial processing evolves from V1 to V4. From the point of view of V1, these families of functions contain elements (those with one index 0 in Cartesian separation) that are very similar to Gabor patches (Swanson, Wilson, and Giese [1984]). From the point of view of V4, the families (in polar separation) contain elements that resemble the non-Cartesian gratings used by Gallant and co-workers. With these stimuli, hypotheses for the organization and relationship of V1 and V4 receptive fields can readily be tested. For example, if indeed V1 neurons are one-dimensional analyzers, there should be little response to functions which, in the Cartesian separation $\psi_{n_x}(x)\psi_{n_y}(y)$, have both indices n_x and n_y nonzero. On the other hand, if V1 neurons only *appear* to be one-dimensional analyzers because they have a limited space-bandwidth product, then neurons that respond to an intrinsically two-dimensional function such as $\psi_1(x)\psi_1(y)$ should be about as common as neurons that respond to a one-dimensional multilobed function, such as $\psi_2(x)\psi_0(y)$, of equal eigenvalue. The hypothesis that processing in V4 is intrinsically two-dimensional, rather than just characterized by an increased space-bandwidth product, makes the converse prediction.

As noted above, the "Gabor" profiles of typical neurons in V1 have only one or two major lobes, and equally well might be replaced by functions such as $\psi_1(x)\psi_0(y)$ and $\psi_2(x)\psi_0(y)$, with an appropriate choice of axes. Wilson [1999], drawing on anatomical work by Schoups, Tootell, and Orban [1995], has suggested that V4 neurons obtain their properties by pooling of V1 responses across multiple orientations, and that circularly symmetric V4 receptive fields might result from uniform pooling of one-dimensional profiles in V1. Appropriately weighted combinations of profiles at orthogonal orientations also can lead to receptive field sensitivity profiles with n-fold rotational symmetry, for arbitrary n, as seen from the relationship between the Cartesian and polar separations described above. Neurons with preferences for n-fold rotational symmetry have been encountered in extrastriate cortices (Tanaka, Saito, Fukada, and Moriya [1991]). It is interesting (and also highly speculative) to note that the relationship between the polar separation and the Cartesian separation provides an economical way to construct receptive fields of arbitrary rotational symmetry.

Our crude analysis (Figure 2.8) of simulated neurons based on the eigenfunctions in polar separation indicates that these receptive field profiles indeed can be tuned for Cartesian and non-Cartesian gratings, as Gallant, Braun, and Van Essen [1993]; Gallant, Connor, Rakshit, Lewis, and Van Essen [1996] found for real V4 neurons. Given the highly schematized and simple nature of our simulation, a detailed correspondence with physiology would be unexpected. Indeed, only certain aspects of Gallant et al.'s results are accounted for by this scheme. Gallant et al. found tuning to hyperbolic gratings that was more prominent than our simulations sug-

gest. They also found that responses were relatively insensitive to translations of the stimulus. This suggests that the very simple scheme we used here (inner product followed by a quadrature-pair calculation) needs to be augmented by additional stages of processing, perhaps in the form of receptive field subunits. More detailed modeling in this direction requires additional physiologic data, and is beyond the scope of this work. However, it is hoped that the considerations presented here will serve as a basis for a sound theoretically-motivated experimental approach to understanding visual processing beyond the earliest cortical stages.

4 Appendix

Normalization of the Eigenfunctions in the Gaussian Case. The normalization of the eigenfunctions (2.29) requires integrals similar to equation (2.58), but with respect to Gaussians of variance other than unity. The generating-function strategy used for the evaluation of the integrals of equation (2.58) readily extends to this situation. First, consider

$$\frac{1}{2\pi V}\iint \sum_0^\infty \frac{z^k t^l\, q_{k,l}(x,y)}{k!\,l!}\,\frac{z'^{k'} t'^{l'}\,\overline{q}_{k',l'}(x,y)}{k'!\,l'!}\exp\Big[-\frac{1}{2V}(x^2+y^2)\Big]\,dx\,dy$$
$$= \frac{1}{2\pi V}\iint Q(z,t)\overline{Q}(x',t')\exp\Big[-\frac{1}{2V}(x^2+y^2)\Big]\,dx\,dy \tag{4.1}$$

in which the sum on the left-hand side is over all values of k, l, k' and l'. Proceeding in the same manner as with equation (2.58), we find

$$\frac{1}{2\pi V}\iint \sum_0^\infty \frac{z^k t^l q_{k,l}(x,y)}{k!\,l!}\,\frac{z'^{k'} t'^{l'}\overline{q}_{k',l'}(x,y)}{k'!\,l'!}\exp\big[\frac{1}{2V}(x^2+y^2)\big]\,dx\,dy$$
$$= \exp\big[2V(zz'+tt')\big]\exp\big[2(V-1)(zt+z't')\,.\big] \tag{4.2}$$

Because of the polar symmetry, terms for which $k-l \neq k'-l'$ are necessarily 0. Extracting and equating terms corresponding to $z^{\mu+\nu}t^\nu(z')^{\mu+\omega}(t')^\omega$ leads to

$$\frac{1}{2\pi V}\iint q_{\mu+\nu,\nu}(x,y)\,\overline{q}_{\mu+\omega,\omega}(x,y)\,\exp\Big[-\frac{1}{2V}(x^2+y^2)\Big]\,dx\,dy$$
$$= (\mu+\nu)!\nu!(\mu+\omega)!\omega!\sum_{\alpha,\beta,\gamma,\delta}\frac{2^{\alpha+\beta+\gamma+\delta}(V-1)^{\alpha+\beta}V^{\gamma+\delta}}{\alpha!\,\beta!\,\gamma!\,\delta!}\,, \tag{4.3}$$

where the sum is over all indices α, β, γ and δ satisfying

$$\alpha+\gamma = \mu+\nu$$
$$\alpha+\delta = \nu$$

$$\begin{aligned}\beta + \gamma &= \mu + \omega \\ \beta + \delta &= \omega\,. \end{aligned} \tag{4.4}$$

Equation (4.3) leads to

$$\begin{aligned}&\frac{1}{2\pi V}\iint q_{\mu+\nu,\nu}(x,y)\,\overline{q}_{\mu+\omega,\omega}(x,y)\,\exp\Big[-\frac{1}{2V}(x^2+y^2)\Big]\,dx\,dy \\ &\qquad = 2^{\mu+\nu+\omega}(\mu+\nu)!\,\nu!\,(\mu+\omega)!\,\omega!\sum_{s}^{\min(\nu,\omega)}\frac{(V-1)^{\nu+\omega-2s}\,V^{\mu+2s}}{(\nu-s)!\,(\omega-s)!\,s!\,(\mu+s)!}\,.\end{aligned} \tag{4.5}$$

The general result (4.5) can also be recast as

$$\begin{aligned}&\frac{1}{V}\int_0^\infty P_{\mu,\nu}(R^2)\,P_{\mu,\omega}(R^2)\,R^{2\mu+1}\,\exp\Big(-\frac{R^2}{2V}\Big)\,dR \\ &\qquad = 2^{\mu+\nu+\omega}(\mu+\nu)!\,\nu!\,(\mu+\omega)!\,\omega!\sum_{s}^{\min(\nu,\omega)}\frac{(V-1)^{\nu+\omega-2s}V^{\mu+2s}}{(\nu-s)!\,(\omega-s)!\,s!\,(\mu+s)!}\,.\end{aligned} \tag{4.6}$$

To see the relationship to the case $V = 1$ of equation (2.58), note that the only nonzero contributions to the sum in equation (4.3) are for $\alpha = \beta = 0$ which forces $\gamma = \mu + \nu = \mu + \omega$ and $\delta = \nu = \omega$. In equations (4.5) and (4.6), this corresponds to the sole term $s = \nu = \omega$, which establishes the reduction to equation (2.60) with $\kappa = \mu + \nu$ and $l = \nu$ and to equation (2.62).

Modest Non-Orthogonality After Shifts. The extent to which shifting a function spoils the orthogonality can be captured by considering the Hermite-function limit (large c) of the one-dimensional eigenfunctions. These eigenfunctions (see equations (2.21) and (2.27)) are given by

$$\varphi_n(x) = h_n(kx)\,\exp\big(-\tfrac{1}{4}k^2x^2\big)\,, \tag{4.7}$$

from which it follows that

$$\sum_{n=0}^{\infty}\frac{z^n}{n!}\varphi_n(x) = \exp\big(-\tfrac{1}{4}k^2x^2 + kxz - \tfrac{1}{2}z^2\big)\,. \tag{4.8}$$

To determine the non-orthogonality of the mth and nth eigenfunctions after positional shifts s, define

$$a_{m,n}(x) = \frac{k}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\varphi_m(x-s)\,\varphi_n(x+s)\,dx\,. \tag{4.9}$$

Consider a generating function defined by

$$A(s,y,z) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{y^m}{m!}\,\frac{z^n}{n!}\,a_{m,n}(s)\,. \tag{4.10}$$

From the generating function (4.8) for φ, it follows that

$$A(s,y,z) = \frac{k}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\bigl(-\tfrac{1}{4}k^2(x-s)^2 + k(x-s)y - \tfrac{1}{2}y^2\bigr) \exp\bigl(-\tfrac{1}{4}k^2(x+s)^2 + k(x+s)z - \tfrac{1}{2}z^2\bigr)\, dx\,, \tag{4.11}$$

and hence,

$$\begin{aligned} A(s,y,z) &= \frac{k}{\sqrt{2\pi}} \exp\bigl(-\tfrac{1}{2}k^2s^2 + ks(-y+z)\bigr) \int_{-\infty}^{\infty} \exp\bigl(-\tfrac{1}{2}k^2x^2 + kx(y+z) - \tfrac{1}{2}y^2 + z^2\bigr)\, dx \\ &= \exp\bigl(-\tfrac{1}{2}k^2s^2 + ks(-y+z) + yz\bigr)\,. \end{aligned} \tag{4.12}$$

Equating coefficients of terms $y^m z^n$ in equations (4.10) and (4.12) yields:

$$a_{m,n}(s) = \sum_{j=0}^{\min(m,n)} (-1)^{m-j} j! \binom{m}{j} \binom{n}{j} (ks)^{m+n-2j} \exp\bigl(-\tfrac{1}{2}k^2s^2\bigr)\,. \tag{4.13}$$

Note that when $s = 0$, the only nonzero term on the right-hand side of equation (4.13) has $m + n - 2j = 0$, which can only occur if $m = n$ (thus, recovering orthogonality when there is no shift). For nonzero s, the values of $a_{m,n}(s)$ are controlled by the Gaussian envelope.

Acknowledgments: This work was supported by NIH NEI EY9314 and DARPA MDA972-01-1-0028. The authors thank Partha Mitra and Larry Sirovich for several helpful discussions, Stan Klein and Hugh Wilson for insightful comments on the manuscript, and Adam Kahn for programming assistance.

References

Abramowitz, M. and I. A. Stegun [1964], *Handbook of Mathematical Functions.* National Bureau of Standards. Reprinted by Dover, New York, 1970.

Atick, J. and A. Redlich [1990], Towards a theory of early visual processing, *Neural Comput.* **2**, 308–320.

Daugman, J. G. [1985], Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, *J Opt Soc Am [A].* **2**, 1160–9.

Erdelyi, A. [1955], *Higher Transcendental Functions.* Bateman Manuscript Project, Vol. II. New York: McGraw-Hill.

Field, D. J. (1994), What is the goal of sensory coding? *Neural Comput.* **6**: 559–601.

Flammer, C. [1957], *Spheroidal Wave Functions.* Stanford: Stanford University Press.

Gabor, D. [1946], Theory of communication, *J. Inst. Elect. Eng.* **93**, 429–457.

Gallant, J. L., J. Braun, and D. C. Van Essen [1993], Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex, *Science*, **259**, 100–3.

Gallant, J. L., C. E. Connor, S. Rakshit, J. W. Lewis, and D. C. Van Essen [1996], Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey, *J Neurophysiol.* **76**: 2718–39.

Grosof, D. H., R. M. Shapley, and M. J. Hawken [1993], Macaque V1 neurons can signal 'illusory' contours, *Nature.* **365**, 550–2.

Halmos, P. [1942], *Finite Dimensional Vector Spaces.* Princeton: Princeton University Press.

Jones, J. P. and L. A. Palmer [1987], An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex, *J Neurophysiol.* **58**, 1233–58.

Klein, S. A. and B. Beutter [1992], Minimizing and maximizing the joint space-spatial frequency uncertainty of Gabor-like functions: comment, *J. Opt. Soc. Am. A* **9**, 337–40.

Knight, B. W. and L. Sirovich [1982], The Wigner transform and some exact properties of linear operators, *SIAM J. Appl. Math.* **42**, 378–389.

Knight, B. W. and L. Sirovich [1986], The eigenfunction problem in higher dimensions: Exact results, *Proc. Natl. Acad. Sci. USA* **83**, 527–530.

Kobatake, E. and K. Tanaka, [1994], Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex, *J Neurophysiol.* **71**, 856–67.

Laughlin, S. B., R. R. de Ruyter van Steveninck, and J. C. Anderson [1998], The metabolic cost of neural information, *Nat Neurosci.* **1**, 36–41.

Levitt, J. B., D. C. Kiper, and J. A. Movshon[1994], Receptive fields and functional architecture of macaque V2, *J Neurophysiol.* **71**, 2517–42.

Marcelja, S. [1980], Mathematical description of the responses of simple cortical cells, *J Opt Soc Am.* **70**, 1297–1300.

Merigan, W. H. [1996], Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques, *Vis Neurosci.* **13**, 51–60.

Percival, D. and A. Walden [1993], *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques.* Cambridge: Cambridge University Press.

Purpura, K. P., J. D. Victor, and E. Katz [1994], Striate cortex extracts higher-order spatial correlations from visual textures, *Proc Natl Acad Sci U S A.* **91**, 8482–6.

Ringach, D. L., G. Sapiro, and R. Shapley [1997], A subspace reverse-correlation technique for the study of visual neurons, *Vision Res.* **37**, 2455–2464.

Rota, G.-C. and B. D. Taylor [1994] The classical umbral calculus, *SIAM J. Math. Anal.* **25**, 694–711.

Rubin, N. [2001], The role of junctions in surface completion and contour matching, *Perception* **30**, 339–366

Schoups, A. A., R. B. Tootell, W. Vanduffel, and G. A. Orban [1995], Use of the double-label deoxyglucose approach to map the orientation columnar system beyond area V1 and V2 in the macaque, and its plasticity, *Soc. Neurosci. Abstr.* **21**, 18.

Sirovich, L. and B. W. Knight [1981], On the eigentheory of operators which exhibit a slow variation, *Quart. Appl. Math.* **38**, 469–488.

Sirovich, L. and B. W. Knight [1982], Contributions to the eigenvalue problem for slowly varying operators, *SIAM J. Appl. Math.* **42**, 356–377.

Sirovich, L. and B. W. Knight [1985], The eigenfunction problem in higher dimensions: Asymptotic Theory, *Proc. Natl. Acad. Sci. USA.* **82**, 8275–8278.

Slepian, D. and H. O. Pollack [1961], Prolate spheroidal wave functions, Fourier analysis and uncertainty I. *Bell Syst. Tech. J.* **40**, 43–64.

Slepian, D. [1964], Prolate spheroidal wave functions, Fourier analysis and uncertainty IV: Extensions to many dimensions; generalized prolate spheroidal functions, *Bell Syst. Tech. J.* **43**, 3009–3057.

Stork, D. G. and H. R. Wilson [1990], Do Gabor functions provide appropriate descriptions of visual cortical receptive fields?, *J. Opt. Soc. Am. A* **7**, 1362–73.

Swanson, W. H., H. R. Wilson, and S. C. Giese [1984], Contrast matching data predicted from contrast increment thresholds, *Vision Res.* **24**, 63–75.

Tanaka, K., H. Saito, Y. Fukada, and M. Moriya [1991], Coding visual images of objects in the inferotemporal cortex of the macaque monkey, *J Neurophysiol.* **66**, 170–89.

Victor, J. D. and S. Brodie [1978], Discriminable textures with identical Buffon needle statistics, *Biological Cybernetics* **31**, 231–234.

Vilenkin, N. J. [1968], *Special Functions and the Theory of Group Representations*. American Mathematical Society.

von der Heydt, R., E. Peterhans, and G. Baumgartner [1984], Illusory contours and cortical neuron responses, *Science* **224**, 1260–2.

Wilkinson, F., H. R. Wilson, and C. Habak [1998], Detection and recognition of radial frequency patterns, *Vision Res.* **38**, 3555–68.

Wilson, H. R. [1999] Non-Fourier cortical processes in texture, form, and motion perception. In *Cerebral Cortex, vol. 13*, pages 445–477. ed: Ulinski et al. Kluwer Academic, New York.

Wilson, H. R. and F. Wilkinson [1998], Detection of global structure in Glass patterns: implications for form vision, *Vision Res.* **38**, 2933–47.

Wolfe, J. M., A. Yee, and S. R. Friedman-Hill, [1992], Curvature is a basic feature for visual search tasks, *Perception* **21**, 465–80.

Xu, Y., S. Haykin, and R. J. Racine [1999], Multiple window time-frequency distribution and coherence of EEG using Slepian sequences and Hermite functions, *IEEE Trans. Biomed. Engrg.* **46**, 861–866.