Published in final edited form as:

J Opt Soc Am A Opt Image Sci Vis. 2009 September ; 26(9): 2067–2092.

Simple Cell Response Properties Imply Receptive Field Structure: Balanced Gabor and/or Bandlimited Field Functions

Davis Cope¹, Barbara Blakeslee^{2,3}, and Mark E. McCourt^{2,3}

¹Department of Mathematics North Dakota State University Fargo, ND 58108-6050

²Department of Psychology North Dakota State University Fargo, ND 58108-6050

³Center for Visual Neuroscience North Dakota State University Fargo, ND 58108-6050

Abstract

The classical receptive fields of simple cells in mammalian primary visual cortex demonstrate three cardinal response properties: 1) they do not respond to stimuli which are spatially homogeneous; 2) they respond best to stimuli in a preferred orientation (direction); and 3) they do not respond to stimuli in other, non-preferred orientations (directions). We refer to these as the Balanced Field Property, the Maximum Response Direction Property, and the Zero Response Direction Property, respectively. These empirically-determined response properties are used to derive a complete characterization of elementary receptive field functions (of cosine- and mixed-type) defined as products of a circularly symmetric weight function and a simple periodic carrier. Two disjoint classes of elementary receptive field functions result: the balanced Gabor class, a generalization of the traditional Gabor filter, and a bandlimited class whose Fourier transforms have compact support (i.e., are zero-valued outside of a bounded range). The detailed specification of these two classes of receptive field functions from empirically-based postulates may prove useful to neurophysiologists seeking to test alternative theories of simple cell receptive field structure, and to computational neuroscientists seeking basis functions with which to model human vision.

Keywords

Simple cells; receptive fields; visual cortex; Gabor filter; Bessel function; bandlimited; Hankel transform; Fourier analysis

1. Introduction

The pioneering work of Hubel and Weisel (1968; 1977) established that neurons in the mammalian primary visual cortex possessed receptive fields that, unlike the circularly concentric receptive fields of subcortical visual neurons, were spatially oriented. The receptive field of one type of cortical neuron, the simple cell, possessed discrete elongated regions within which stimulation by light evoked either excitation or inhibition. This geometry suggested that simple cells might be especially responsive to visual stimuli consisting of bars or edges. Reinterpreting the presumptive role of simple cells as bar and edge detectors, DeValois and colleagues (1978; 1988) demonstrated that the periodic

Corresponding author: Mark E. McCourt, Center for Visual Neuroscience, Department of Psychology, North Dakota State University, Fargo, ND 58108-6050. mark.mccourt@ndsu.edu. Mark E. McCourt Department of Psychology NDSU Dept. 2765 PO Box 6050 Fargo, ND 58108-6050 mark.mccourt@ndsu.edu 701.231-8625.

Commercial relationships: none.

structure of simple cell receptive fields rendered them highly selective for stimulus spatial frequency and suggested that they might therefore constitute the basis functions for a piecewise Fourier (or wavelet) encoding of the retinal intensity distribution. Much attention was concurrently devoted to developing a mathematical description of the simple cell receptive field function. The first such descriptions, formulated by Marcelja (1980) and Daugman (1980; 1985), inspired by the earlier work of Gabor (1946), showed that the spatial sensitivity of simple cells was well described by a two-dimensional Gaussian-damped sinewave carrier signal, a so-called Gabor function. Due to some undesirable features of the Gabor function, in particular the fact that it does not integrate to zero for all carrier phases, a variety of competing mathematical formulations have been proposed, which include the difference-of-Gaussians (Heggelund, 1986a; b), the Laplacian-of-a-Gaussian (Marr & Hildreth, 1980), the log-Gabor (Field, 1987) and the Cauchy function (Klein & Levi, 1985) - see Wallis (2001) for a critical review. However, the Gabor model has withstood the test of repeated physiological (Kulikowski & Bishop, 1981; Jones & Palmer, 1987; DeValois & DeValois, 1988; Ringach, 2002), and psychophysical (Watson, Barlow & Robson, 1983; Foley, Varadharajan, Koh & Farias, 2007) verification, and has therefore emerged as the most commonly accepted mathematical description of cortical simple cell receptive field structure. Theorists interested in developing mathematical descriptions and computational models of various aspects of human spatial vision therefore frequently employ Gabor filters as basis functions (Heitger, Rosenthaler, von der Heydt, Peterhans & Kubler, 1992; Lee, 1996; Potzsch, Kruger & von der Malsburg, 1996; Petkov, 1995; Petkov & Kruzinga, 1997).

Experimental studies have established that simple cells show a variety of well-defined behaviors (Ringach, 2002). We are concerned here with three essential properties of simple cell receptive fields: 1) the *Balanced Field Property* (i.e., spatially homogeneous patterns produce a zero response); 2) the *Zero Response Direction Property* (i.e., there is a direction (viz, stimulus orientation) which elicits a zero response to sinusoidal gratings); and 3) the *Maximum Response Direction Property* (the maximum response decreases monotonically to zero as direction (i.e, stimulus orientation) changes from the optimal direction to a direction perpendicular to the optimal). Note that we use the term “direction” to describe what vision scientists commonly refer to as “orientation”, where grating direction is orthogonal to grating orientation.

In this paper we use these three empirical response properties of simple cells as postulates, which are themselves independent of any particular receptive field structure, and apply them to obtain a complete classification of elementary receptive field functions (of cosine- and mixed-type), which are defined as the product of a circularly symmetric weight function and a simple periodic carrier. Thus, we describe all possible types of elementary receptive field functions which satisfy these empirical constraints. Two disjoint classes of receptive field functions result. The first class, designated here as balanced Gabor fields, is found to be a natural generalization of the traditional Gabor receptive field model. It includes as its simplest case the simple balanced Gabor. The simple balanced Gabor has a receptive field structure similar to the traditional Gabor receptive field model, but integrates to zero for all spatial phases of the periodic carrier function. Balanced Gabor filters possess all of the desirable features of the traditional Gabor model, viz., spatial frequency and orientation (direction) tuning, etc., but correct what is commonly regarded as the chief deficiency of the traditional Gabor, which is that it possesses a non-zero DC response. The second class, designated here as bandlimited fields (Victor & Knight, 2003) also possess well-behaved spatial frequency and orientation response properties, but differ sufficiently from balanced Gabor functions in other respects that an empirical determination of which class best describes simple cell receptive field structure may be possible.

2. Formalization of postulated simple cell receptive field properties

This section states the Balanced Field, Zero Response Direction (ZRD), and Maximum Response Direction (MRD) Properties and provides a concise mathematical formulation. It should be noted that these three properties are cumulative, not independent, that is, the Maximum Response Direction (MRD) Property mathematically implies the Zero Response Direction (ZRD) Property, which in turn implies the Balanced Field Property. Although experimental procedure may emphasize the maximum response direction, the zero response direction has more immediate theoretical implications, and this cumulative formulation has been found useful for the development of such implications. It is assumed that the integral (2.1) describes the interaction of the receptive field with the stimulus, but no assumption about a specific structure of the receptive field is made in this section. Section 3 will make such an assumption in defining an *elementary receptive field function* and a mathematically rigorous formulation will be given there.

We define the *visual field* as a plane, described by the cartesian coordinates (x_1, x_2) (briefly, $\vec{x}$). We define visual stimuli, or *patterns* in the visual field, as nonnegative, bounded, and

possibly time-dependent functions $p(\vec{x}; t)$ on the plane. The *receptive field* of the simple cell is described with respect to a fixed reference location, its *center*, taken here to be the

origin. The receptive field is modeled by the *receptive field function* $R(\vec{x})$, an absolutely integrable function which is assumed to interact with the stimulus pattern by

$$R \circ p(t) = \int_{\mathbb{R} \times \mathbb{R}} R(\vec{x}) p(\vec{x}; t) d\vec{x} \quad (2.1)$$

Here $R \circ p$ is the *response of the receptive field*, briefly, the *response*. The response to a spatially homogeneous pattern $p = c$ is $c R_0$, where

$$R_0 := \int_{\mathbb{R} \times \mathbb{R}} R(\vec{x}) d\vec{x} \quad (2.2)$$

The properties below describe the response $R \circ p$ to a *sinusoidal grating stimulus* p , which (normalized to mean value 1) is a pattern of the form

$$p(\vec{x}) := 1 + c_p \cos \left(2\pi \frac{\vec{d}(\alpha_p) \cdot \vec{x}}{\lambda_p} - \phi_p \right) \quad (2.3)$$

where c_p is the *grating contrast* ($-1 \leq c_p \leq +1$), $\lambda_p > 0$ is the *grating wavelength*, $\vec{d}$ is a unit

vector defining *grating direction* $\vec{d}(\alpha_p) \cdot \vec{x} = \cos(\alpha_p)x_1 + \sin(\alpha_p)x_2$, and ϕ_p is a *phase shift*, which is time-dependent in the case of a drifting grating. The subscript “P” designates pattern parameters. The pattern is a periodic sequence of bright and dark bands

perpendicular to $\vec{d}$, normalized to have mean value one. Notice the grating parameter sets $c_p, \lambda_p, \alpha_p, \phi_p$ and $-c_p, \lambda_p, \alpha_p, \phi_p \pm \pi$ and $c_p, \lambda_p, \alpha_p \pm \pi, -\phi_p$ all describe the same grating and hence will produce the same response.

The response to a sinusoidal grating stimulus is closely related to the Fourier transform of the receptive field function. Specifically, the Fourier transform $F_R(s_1, s_2)$ of $R(x_1, x_2)$ is defined by

$$F_R(s_1, s_2) := \int_{\mathbb{R} \times \mathbb{R}} \exp(-2\pi i(s_1 x_1 + s_2 x_2)) R(x_1, x_2) dx_1 dx_2 \quad (2.4)$$

The response $c R_0$ to the constant pattern is related to the Fourier transform by

$$R_0 = F_R(0, 0) \quad (2.5)$$

while the *response function*, that is, the response $R \circ p$ to a sinusoidal grating pattern p , is given by (since R is real-valued)

$$\begin{aligned} R \circ p &= \int_{\mathbb{R} \times \mathbb{R}} R(\vec{x}) \left(1 + c_p \cos \left(2\pi \frac{\vec{d}(\alpha_p) \cdot \vec{x}}{\lambda_p} - \phi_p \right) \right) d\vec{x} \\ &= R_0 + c_p \operatorname{Re} \left[\oplus^{i\phi_p} F_R \left(\frac{1}{\lambda_p} \cos(\alpha_p), \frac{1}{\lambda_p} \sin(\alpha_p) \right) \right] \end{aligned} \quad (2.6)$$

which determines the transform (e.g. $\operatorname{Re}[F_R]$, $-\operatorname{Im}[F_R]$ for $\phi_p = 0, \pi/2$). The *max-response* M to a sinusoidal grating stimulus is the maximum over all phases ϕ_p :

$$\begin{aligned} M(c_p, \alpha_p, \lambda_p) &:= \max_{\phi_p} (R \circ p) \\ &= \max_{\phi_p} \left(R_0 + c_p \operatorname{Re} \left[\oplus^{i\phi_p} F_R \left(\frac{1}{\lambda_p} \cos(\alpha_p), \frac{1}{\lambda_p} \sin(\alpha_p) \right) \right] \right) \\ &= R_0 + |c_p| \left| F_R \left(\frac{1}{\lambda_p} \cos(\alpha_p), \frac{1}{\lambda_p} \sin(\alpha_p) \right) \right| \end{aligned} \quad (2.7)$$

Since R is real-valued, the Fourier transform satisfies $\bar{F}_R(s_1, s_2) = F_R(-s_1, -s_2)$. Consequently, $M(c_p, \alpha_p, \lambda_p) = M(|c_p|, \alpha_p, \lambda_p) = M(|c_p|, \alpha_p \pm \pi, \lambda_p)$. In particular, M has period π with respect to α_p .

Since the response function, that is, the response over all grating directions and wavelengths, determines the Fourier transform of R , sinusoidal grating experiments have fundamental significance for the study of the receptive field. If equation (2.1) completely described the response to a visual stimulus, then the the Fourier transform, that is, the response function, would completely determine the receptive field. In practice, the response necessarily includes nonlinear behavior not described by equation (2.1), but it is widely accepted that the equation models a major portion of the response, and the response function determined by sinusoidal grating experiments plays a correspondingly major role. A second reason for the fundamental significance of grating experiments is that simple cells respond in a remarkably well-defined way to sinusoidal grating stimuli, showing a maximum response for an optimum grating direction (orientation) and optimum grating wavelength, then decreasing steadily as these parameters vary from the optimum with, in particular, a zero response for grating directions perpendicular to the optimum direction (orientation). Our postulates are expressed in terms of such observed behavior for the response to sinusoidal gratings.

The first postulated property of simple cells is that spatially homogeneous patterns produce a zero response, *i.e.*, the

Balanced Field Property

The response $R \circ p$ to a spatially homogeneous pattern $p \equiv \text{constant}$ is zero.

Such patterns correspond to a zero-contrast grating. The constant scales out, and the mathematical form is

$$R_0 = F_R(0, 0) = 0 \quad (2.8)$$

The second postulated property of simple cells is that there is a stimulus direction (*i.e.*, grating orientation) which elicits a zero response, the *Zero Response Direction Property*:

Zero Response Direction (ZRD) Property (Form 1)

There exist values α_{ZR} and $c_{ZR} > 0$ such that $R \circ p = 0$ for all sinusoidal gratings p with grating direction α_{ZR} and contrast c_{ZR} .

The mathematical form is that, for all grating phases ϕ_P and wavelengths $\lambda_P > 0$,

$$R_0 + c_{ZR} \operatorname{Re} \left[e^{i\phi_P} F_R \left(\frac{1}{\lambda_P} \cos(\alpha_{ZR}), \frac{1}{\lambda_P} \sin(\alpha_{ZR}) \right) \right] = 0 \quad (2.9)$$

Notice the ZRD Property implies the Balanced Field Property (consider the phase shifts ϕ_P , $\phi_P + \pi$ and add), that is, $R_0 = 0$. Since (2.9) can then hold for all ϕ_P if and only if $F_R = 0$, the ZRD Property is equivalent to the following formulation, which drops all reference to grating contrast and phase:

Zero Response Direction (ZRD) Property (Form 2)

There exists a value α_{ZR} such that, for all $\lambda_P > 0$,

$$F_R \left(\frac{1}{\lambda_P} \cos(\alpha_{ZR}), \frac{1}{\lambda_P} \sin(\alpha_{ZR}) \right) = 0 \quad (2.10)$$

To see that Form 2 implies Form 1, let $\lambda_P \rightarrow \infty$ in (2.10) to show the field is balanced, then (2.9) follows for arbitrary contrast. The direction α_{ZR} is a *zero response direction* and is only determined up to an additive multiple of π ; values differing by multiples of π are *equivalent directions*, and multiple, that is, nonequivalent, zero response directions (orientations) are allowed.

Incidentally, in modeling the neurophysical response to excitation and inhibition, the response is sometimes taken to be the positive part of the calculated numerical value since firing rates cannot be negative. Using that convention, a zero response would mean $\operatorname{Pos}(R \circ p) = 0$, that is, $R \circ p \leq 0$. That convention is not used here: *response* means $R \circ p$, and a *zero response* means $R \circ p = 0$.

The third postulated property is essentially that the maximum response decreases monotonically to zero as direction (*i.e.*, stimulus orientation) changes from the optimal direction to a direction perpendicular to the optimal, the *Maximum Response Direction Property*:

Maximum Response Direction (MRD) Property (Form 1)

There exists a value $c_{MR} > 0$ such that $M(c_{MR}, \alpha_P, \lambda_P)$ is not identically zero and a value α_{MR} such that, for all sinusoidal gratings p with contrast c_{MR} and at each fixed wavelength λ_P , either the response $M(c_{MR}, \alpha_P, \lambda_P)$ is zero for all grating directions α_P or it is strictly decreasing to zero (and remains zero) as the grating direction changes from α_{MR} .

The range $|\alpha_P - \alpha_{MR}| \leq \pi/2$ covers one period of $M(c_{MR}, \alpha_P, \lambda_P)$, and the MRD Property implies $M(c_{MR}, \alpha_{MR} \pm \pi/2, \lambda_P) = 0$ for each wavelength λ_P . Thus, the MRD Property implies the ZRD Property with $\alpha_{ZR} = \alpha_{MR} \pm \pi/2$. The Balanced Field Property therefore

holds as well, so that $M(c_{MR}, \alpha_p, \lambda_p) = |c_{MR}| \left| F_R \left(\frac{1}{\lambda_p} \cos(\alpha_p), \frac{1}{\lambda_p} \sin(\alpha_p) \right) \right|$. A further implication is that ZRDs necessarily occur as a sector (possibly reducing to a single line). For example, if $\alpha_{ZR} = \alpha_{MR} + \zeta_0$ with $0 < \zeta_0 < \pi/2$ is a ZRD, then, for each fixed λ_p , the response will have dropped to zero when the stimulus direction α_p reaches α_{ZR} and then remains zero as α_p increases. The corresponding range $\alpha_{MR} + \zeta_0$ to $\alpha_{MR} + \pi/2$ is then a sector of ZRDs.

The MRD Property is equivalent to the following formulation, which drops all reference to grating contrast and phase:

Maximum Response Direction (MRD) Property (Form 2)

$F_R \left(\frac{1}{\lambda_p} \cos(\alpha_p), \frac{1}{\lambda_p} \sin(\alpha_p) \right) \neq 0$ for some pair α_p, λ_p and there exists a value α_{MR} such that,

at each fixed wavelength λ_p , either $\left| F_R \left(\frac{1}{\lambda_p} \cos(\alpha_p), \frac{1}{\lambda_p} \sin(\alpha_p) \right) \right|$ is zero for all grating directions α_p or it is strictly decreasing to zero (and remains zero) as $|\alpha_p - \alpha_{MR}|$ increases.

To see that Form 2 implies Form 1, notice that $\left| F_R \left(\frac{1}{\lambda_p} \cos(\alpha_p), \frac{1}{\lambda_p} \sin(\alpha_p) \right) \right|$ has period π with respect to α_p . Then $|\alpha_p - \alpha_{MR}| < \pi/2$ covers one period, the argument above can be repeated to obtain the ZRD and Balanced Field Properties, and the behavior for $M(c_{MR}, \alpha_p, \lambda_p)$ stated in Form 1 follow.

The direction α_{MR} is a *maximum response direction* and is only determined up to an additive multiple of π ; values differing by multiples of π are *equivalent directions (orientations)*. It should be emphasized that this property not only refers to the existence of a direction (orientation) of maximum response but also to the monotonic decrease in response with increasing angular distance from the optimal direction. Such monotonic behavior implies that a maximum response direction, if it exists, is unique.

It should be noted that the properties are robust with respect to nonlinear processing. For example, if the response defined by (2.1) is following by a nonlinear operation, then the properties will continue to hold for the new result given some mild conditions on the nonlinear operation, such as monotonicity and having zero as a fixed point.

3. Elementary receptive field functions and Hankel transforms

Receptive fields are typically modeled as a product of a *weight function*, providing localization of the response, and an oscillatory *carrier function*, providing directionality (orientation) to the response. This section defines an *elementary receptive field function* to be the product of a circularly symmetric weight function and a simple periodic carrier. For such receptive field functions, the response to a sinusoidal grating can be expressed in terms of the Hankel transform of the weight function. This section describes the Hankel transform and its use to provide reduced mathematical formulations of the Zero Response Direction and Maximum Response Direction Properties for elementary receptive field functions. These reduced formulations are the basis for the results of Section 4.

We define an *elementary receptive field function* (centered at the origin) to be the product of a circularly symmetric weight function q and a simple periodic carrier function, that is,

$$R(\vec{x}) := \frac{2\pi}{\lambda_R^2} q\left(2\pi \frac{|\vec{x}|}{\lambda_R}\right) \left(\cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) - b_R \right) \quad (3.1)$$

where $\lambda_R > 0$ is the *spatial wavelength*, α_R is the *field orientation*, ϕ_R is the *phase shift*, and b_R is the *balancing parameter* (the subscript “R” designates receptive field parameters). The weight function $q(r)$ may be negative and is assumed to satisfy the mild regularity conditions

$$q(r) \text{ is } C(0, +\infty), q(r) \text{ is absolutely integrable on } (0, +\infty), \text{ and } \int_0^\infty q(r) r dr = 1 \quad (3.2)$$

It is useful to distinguish the cases $\sin(\phi_R) = 0$, $\cos(\phi_R) \sin(\phi_R) \neq 0$, $\cos(\phi_R) = 0$ as elementary receptive field functions of *cosine-type*, *mixed-type*, and *sine-type*, respectively. Note that a specific form is assumed for the carrier function, but the weight function (q) is only required to be circularly symmetric and is otherwise unconstrained.

Section 2 formulated properties of the receptive field $R(x_1, x_2)$ in terms of the response to sinusoidal gratings, then gave mathematical formulations in terms of the Fourier transform $F_R(s_1, s_2)$. The form (3.1) for the receptive field is essentially a modulated circularly symmetric function. The 2D-Fourier transform of a circularly symmetric function is also a circularly symmetric function, and the relation between these two functions of a single radial variable is given by the *Hankel transform*:

$$\text{If } f(x_1, x_2) = g(r) \text{ where } r = (x_1^2 + x_2^2)^{1/2}, \text{ then } F_f(s_1, s_2) = 2\pi H_g(2\pi\rho) \text{ where } \rho = (s_1^2 + s_2^2)^{1/2} \quad (3.3)$$

where H_g is the Hankel transform of g (and F_f is defined by (2.4)). The transform and inverse transform are given specifically by

$$H_f(\rho) = \int_0^\infty J_0(\rho r) f(r) r dr \quad \text{and} \quad f(r) = \int_0^\infty J_0(\rho r) H_f(\rho) \rho d\rho \quad (3.4)$$

Here r represents distance from the origin, ρ can be thought of as a radial frequency, and $J_0(x)$ is the Bessel function of order 0. Bessel functions are damped oscillatory waveforms which constitute the basis for Hankel analysis/synthesis, just as sinusoids form the basis functions for Fourier analysis/synthesis. In particular, conditions (3.2) imply

$$H_q(\rho) \text{ is } C[0, +\infty), H_q(0) = 1, \text{ and } H_q(+\infty) = 0 \quad (3.5)$$

Since the elementary receptive field functions R given by (3.1) are simple modulations of the circularly symmetric weight functions q , the Fourier transform F_R can be expressed in terms of translates of the Fourier transform F_q which in turn can be expressed in terms of the Hankel transform H_q :

$$F_R(s_1, s_2) = \frac{1}{2} \left(e^{+i\phi_R} H_q \left(\sqrt{(\lambda_R s_1 + \cos(\alpha_R))^2 + (\lambda_R s_2 + \sin(\alpha_R))^2} \right) + e^{-i\phi_R} H_q \left(\sqrt{(\lambda_R s_1 - \cos(\alpha_R))^2 + (\lambda_R s_2 - \sin(\alpha_R))^2} \right) \right) - b_R H_q \left(\lambda_R \sqrt{s_1^2 + s_2^2} \right) \quad (3.6)$$

Section 3 noted general relations between the Fourier transform F_R of the receptive field function R and its response to a sinusoidal grating, and the corresponding quantities for an elementary receptive field function therefore reduce to expressions involving the Hankel transform H_q . Specifically, for constant patterns, (2.5) becomes:

$$R_0 = \cos(\phi_R) H_q(1) - b_R H_q(0) \quad (3.7)$$

The response function $R \circ p$ of the elementary receptive field function to a sinusoidal grating pattern p , given by (2.6), becomes:

$$R \circ p = R_0 + c_p \left(\frac{1}{2} \left(\cos(\phi_p + \phi_R) H_q \left(\sqrt{1 + 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) + \cos(\phi_p - \phi_R) H_q \left(\sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) \right) - b_R \cos(\phi_p) H_q \left(\frac{\lambda_R}{\lambda_p} \right) \right) \quad (3.8)$$

The max-response M of an elementary receptive field function to a sinusoidal grating pattern, given by (2.7), becomes:

$$M(c_p, \alpha_p, \lambda_p) = \max_{\phi_p} (R \circ p) = R_0 + \frac{1}{2} |c_p| N(\alpha_p, \lambda_p) \quad (3.9a)$$

where the *amplitude factor* $N(\alpha_p, \lambda_p) = 2 |F_R(\cos(\alpha_p) / \lambda_p, \sin(\alpha_p) / \lambda_p)|$ can be obtained by expanding (3.8) in terms of $\cos(\phi_p)$ and $\sin(\phi_p)$ and is now given by

$$N(\alpha_p, \lambda_p)^2 = \left(\cos(\phi_R) \left(H_q \left(\sqrt{1 + 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) + H_q \left(\sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) \right) - 2 b_R H_q \left(\frac{\lambda_R}{\lambda_p} \right) \right)^2 + \left(\sin(\phi_R) \left(H_q \left(\sqrt{1 + 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) - H_q \left(\sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) \right) \right)^2 \quad (3.9b)$$

Note that (3.7, 3.8, 3.9) are completely general formulas describing, in terms of the Hankel transform H_q , the response of elementary receptive field functions (3.1) to sinusoidal grating stimuli.

Section 2 formulated the three properties in terms of the Fourier transform F_R of the general receptive field function R . Using (3.6), those results now become formulations of the properties for an elementary receptive field function (3.1) in terms of the Hankel transform H_q .

The first result follows immediately from (3.7) and uniquely determines the balancing parameter b_R :

Lemma 1

Let the receptive field function $R(\vec{x})$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Then $R(\vec{x})$ has the Balanced Field Property if and only if

$$b_R = \cos(\phi_R) H_q(1) \quad (3.10)$$

The second result follows directly from (2.10), from (3.5), and the representation (3.6) with $\alpha_p = \alpha_{ZR}$:

Lemma 2

Let the receptive field function $R(\vec{x})$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Then $R(\vec{x})$ has the ZRD Property if and only if there exists a value α_{ZR} such that the following relations hold for all $\lambda_p > 0$:

$$\cos(\phi_R) \left(H_q \left(\sqrt{1 + 2 \cos(\alpha_{ZR} - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) + H_q \left(\sqrt{1 - 2 \cos(\alpha_{ZR} - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) \right) = 2 b_R H_q \left(\frac{\lambda_R}{\lambda_p} \right) \quad (3.11a)$$

$$\sin(\phi_R) \left(H_q \left(\sqrt{1 + 2 \cos(\alpha_{ZR} - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) + H_q \left(\sqrt{1 - 2 \cos(\alpha_{ZR} - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) \right) = 0 \quad (3.11b)$$

$$b_R = \cos(\phi_R) H_q(1) \quad (3.11c)$$

In such a case, α_{ZR} is a zero response direction.

The third result is a direct restatement of the MRD Property (Form 2) in terms of the amplitude factor $N(\alpha_p, \lambda_p)$:

Lemma 3

Let the receptive field function $R(\vec{x})$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Then $R(\vec{x})$ has the MRD Property if and only if $N(\alpha_p, \lambda_p) = 0$ for some pair α_p, λ_p and there exists a value α_{MR} such that, at each fixed wavelength λ_p , either $N(\alpha_p, \lambda_p)$ is zero for all grating directions α_p or it is strictly decreasing to zero (and remains zero) as $|\alpha_p - \alpha_{MR}|$ increases.

Section 2 observed that an MRD, if it exists, is unique. It is evident from (3.9b) that $N(\alpha_p, \lambda_p)$ is an even function of $\alpha_p - \alpha_R$, which implies $\alpha_{MR} = \alpha_R$. Section 2 also observed that the MRD Property implies the ZRD Property, specifically, that a ZRD is given by $\alpha_{ZR} = \alpha_R \pm \pi/2$.

4. Characterization of elementary receptive field functions with postulated properties

This section first states Theorems A.1 and A.2, which describe elementary receptive field functions with the Zero Response Direction Property, and then states the main result, Theorem B.1, a complete characterization of elementary receptive field functions with the Maximum Response Direction Property for cosine-type and mixed-type fields. Related results for sine-type fields are summarized in the course of discussion as Theorems B.2 and B.3. The proofs for A.1 and A.2 involve a preparatory lemma and are given in Appendix A. The proof for B.1 builds on A.1 and A.2 and is given in Appendix B.

Theorem A.1 (Cosine-type and mixed-type receptive field functions)

Let the receptive field function $R\left(\frac{\cdot}{x}\right)$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Assume $\cos(\phi_R) \neq 0$. Then $R\left(\frac{\cdot}{x}\right)$ has the ZRD Property if and only if $H_q(\rho)$ satisfies one of the following cases:

1. $0 < H_q(1) < 1$ and $H_q(\rho) = H_q(1)\rho^2 R(\rho^2)$ where $R(y)$ is $C[0, \infty)$ with $R(0) = 1$ and $R(y+1) = R(y)$. In this case, there is exactly one ZRD given by $\alpha_{ZR} - \alpha_R = \frac{\pi}{2}$.
2. $-1 < H_q(1) < 0$ and $H_q(\rho) = (-H_q(1))\rho^2 R(\rho^2)$ where $R(y)$ is $C[0, \infty)$ with $R(0) = 1$ and $R(y+1) = -R(y)$. In this case, there is exactly one ZRD given by $\alpha_{ZR} - \alpha_R = \frac{\pi}{2}$.
3. $H_q(1) = 0$ and $H_q(\rho) = 0$ on the interval $0 < \sin(\zeta_R) - \rho < +\infty$ but on no larger interval. In this case, there is a sector of ZRDs and the sector is given by $0 < \zeta_R \leq |\alpha_{ZR} - \alpha_R| \leq \frac{\pi}{2}$.
4. $0 < |H_q(1)| < 1$ and $\cos(\phi_R) = \pm 1$ and, for some $0 < \zeta_0 \leq \frac{\pi}{2}$, $H_q(\rho)$ satisfies

$$H_q\left(\sqrt{1-2\rho\cos(\zeta_0)+\rho^2}\right) + H_q\left(\sqrt{1+2\rho\cos(\zeta_0)+\rho^2}\right) = 2H_q(1)H_q(\rho) \text{ for } \rho \geq 0.$$

In this case, there are ZRDs given by $\alpha_{ZR} - \alpha_R = \pm\zeta_0$.

Theorem A.2 (Sine-type receptive field functions)

Let the receptive field function $R\left(\frac{\cdot}{x}\right)$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Assume $\cos(\phi_R) = 0$. Then:

1. $R\left(\frac{\cdot}{x}\right)$ has a ZRD given by $\alpha_{ZR} - \alpha_R = \frac{\pi}{2}$.
2. $R\left(\frac{\cdot}{x}\right)$ has more than one ZRD if and only if $H_q(\rho) = 0$ on an interval $\sin(\zeta_0) - \rho < +\infty$ with $0 < \sin(\zeta_R) < 1$. In this case, taking $[\sin(\zeta_R), +\infty)$ to be the largest such interval on which $H_q(\rho)$ vanishes, there is a sector of ZRDs and the sector is given by $0 < \zeta_R \leq |\alpha_{ZR} - \alpha_R| \leq \frac{\pi}{2}$.

Note that Theorem A.1(4) is simply a remaining case, where the theorem does not provide a definitive conclusion. We conjecture that no $H_q(x)$ exists, that is, that the case is vacuous, but we have been unable to settle the conjecture under the sole condition that a ZRD exists. This case is eliminated at the next stage, where the MRD Property is assumed.

Theorem B.1

Let the receptive field function $R\left(\frac{\cdot}{x}\right)$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Assume $\cos(\phi_R) \neq 0$. Then $R\left(\frac{\cdot}{x}\right)$ has the MRD Property if and only if $H_q(\rho)$ satisfies one of the following cases:

1. $H_q(\rho) = e^{-c\rho^2} F(\rho^2)$ where $c > 0$, $F(y)$ is continuous and positive on the real line, $F(0) = 1$, $F(y+1) = F(y)$, and

(*) if $\cos(\phi_R) = \pm 1$, then, for each a , the function

$$f(a; y) = e^{cy} F(a-y) + e^{-cy} F(a+y) - 2F(a) \quad (4.1a)$$

is initially zero on $0 \leq y \leq \delta_a$ and strictly increasing on $\delta_a \leq y < +\infty$;

(**) if $\cos(\phi_R) \neq \pm 1$, then, for each a , the function

$$(\cos(\phi_R))^2 (e^{cy} F(a-y) + e^{-cy} F(a+y) - 2F(a))^2 + (\sin(\phi_R))^2 (e^{cy} F(a-y) - e^{-cy} F(a+y))^2 \quad (4.1b)$$

is strictly increasing on $0 \leq y < +\infty$.

In (*), $\delta_a > 0$ (that is, $f(a; y)$ is initially zero on a nontrivial interval) if and only if the value a satisfies

$$e^{-cy} F(y) = e^{-ca} F(a) - K(a)(y-a) \text{ on } |y-a| \leq \delta_a \text{ for some constant } K(a) > 0 \quad (4.1c)$$

For both (*) and (**):

- a. the MRD is given by $\alpha_{MR} = \alpha_R$;
 - b. the ZRD is unique and given by $\alpha_{ZR} = \alpha_R + \frac{\pi}{2}$;
 - c. $e^{-cy} F(y)$ is positive and strictly decreasing for all y .
2. $H_q(\rho)$ is strictly decreasing on $0 \leq \rho < \sin(\zeta_R)$ and $H_q(\rho) = 0$ on $\sin(\zeta_R) \leq \rho < +\infty$. In this case:
 - a. the MRD is given by $\alpha_{MR} = \alpha_R$;
 - b. there is a sector of ZRDs given by $0 < \zeta_R \leq |\alpha_{ZR} - \alpha_R| \leq \frac{\pi}{2}$.

In result (1) of Theorem B.1, expression (4.1a) is the special case of (4.1b) corresponding to $\cos(\phi_R) = \pm 1$. Parts (*) and (**) can be considered as providing tests for candidate periodic functions $F(y)$ in forming the Hankel transform. Notice it is sufficient to test the expressions for $0 \leq a < 1$ since $F(y)$, and thus the given expressions as well, has period 1. However, it is useful to leave the criterion in terms of arbitrary a .

Theorem B.1 completely characterizes elementary receptive field functions $R\left(\frac{\cdot}{x}\right)$ (of cosine-type and mixed-type) possessing the MRD Property in terms of the Hankel transform $H_q(\rho)$ of the field weight functions $q(r)$. For the moment, let us refer to weights $q(r)$ occurring under result (1) as *Type I* weights and those under result (2) as *Type II*. These two classes of weight functions are completely disjoint, since the Hankel transform $H_q(\rho)$ for a Type I weight is positive for all ρ , while the Hankel transform for a Type II weight is zero outside a finite interval. The characterization theorem thus results in two disjoint classes of receptive field functions $R\left(\frac{\cdot}{x}\right)$, each possessing the MRD Property.

Type I and Type II weights yield quite different formulas for the amplitude response factor $N(\alpha_p, \lambda_p) = 0$ given by (3.9b), and the formulas may be helpful in understanding the results

of Theorem B.1. (The formulas continue to hold for sine-type fields.) For Type I weights with Hankel transform $H_q(\rho) = e^{-c\rho^2} F(\rho^2)$ where $F(y)$ has period one, the factor is

$$N(\alpha_p, \lambda_p)^2 = \exp\left(-c\left(1 + \frac{\lambda_R^2}{\lambda_p^2}\right)\right) \cdot T_F\left(\phi_R, c, 1 + \frac{\lambda_R^2}{\lambda_p^2}; 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) \quad (4.2a)$$

where

$$T_F(\phi_R, c, a; y) := (\cos(\phi_R))^2 (e^{cy} F(a-y) + e^{-cy} F(a+y) - 2F(a))^2 + (\sin(\phi_R))^2 (e^{cy} F(a-y) - e^{-cy} F(a+y))^2 \quad (4.2b)$$

Comparison of (4.2) with result (1) shows that (*) or (**) imply that, for fixed λ_p , $N(\alpha_p, \lambda_p)$ is strictly decreasing as $|\alpha_p - \alpha_R|$ increases. For Type II weights with Hankel transform $H_q(\rho)$ vanishing on some interval $\sin(\zeta_R) \rho < +\infty$, the factor is simply

$$N(\alpha_p, \lambda_p) = H_q\left(\sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}}\right) \quad (4.3)$$

and the result (2) condition that $H_q(\rho)$ is strictly decreasing implies that $N(\alpha_p, \lambda_p)$ is strictly decreasing as $|\alpha_p - \alpha_R|$ increases.

Type II weights are well-defined by result (2), in particular, the definition is independent of the receptive field parameters. In contrast, the definition of Type I weights in result (1) depends on the receptive field phase ϕ_R . That is, $q(r)$ is a Type I weight if its Hankel transform $H_q(\rho)$ satisfies the conditions of result (1) for *some* ϕ_R , and it may be that $q(r)$ satisfies the conditions (*) or (**) for some values of ϕ_R and not for others. Roughly speaking, test (**) becomes less restrictive as $\cos(\phi_R)$ approaches 0. The following theorem illustrates this idea by showing that cosine-type receptive field functions can be used to construct receptive field functions for both mixed-type and sine-type fields. This theorem is the basis for the definition of *balanced Gabor weights* in Section 5a, the subclass of Type I weights that can be used to define elementary receptive field functions with the MRD Property for arbitrary field phase. Type II weights are precisely the *bandlimited weights* of Section 5b. Notice Theorem B.2 yields sine-type receptive field functions.

Theorem B.2

Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Assume:

- a. $H_q(\rho) = e^{-c\rho^2} F(\rho^2)$ where $c > 0$, $F(y)$ is continuous and positive on the real line, $F(0) = 1$, $F(y+1) = F(y)$.
- b. $e^{-cy} F(y)$ is positive and strictly decreasing for all y .
- c. For each a , the function $f(a; y) = e^{cy} F(a-y) + e^{-cy} F(a+y) - 2F(a)$ is initially zero on $0 \leq y \leq \delta_a$ and strictly increasing on $\delta_a \leq y < +\infty$.
- d. In (c), $\delta_a > 0$ if and only if the value a satisfies $e^{-cy} F(y) = e^{-ca} F(a) - K(a)(y-a)$ on $|y-a| \leq \delta_a$ for some constant $K(a) > 0$.

Then the elementary receptive field functions $R\left(\frac{\cdot}{x}\right)$ defined by (3.1) with $b_R = \cos(\phi_R) e^{-c}$ have the MRD Property for all values of the field parameters, in particular, all values of the field phase ϕ_R .

Theorem B.2 follows directly from result (1) of Theorem B.1:

- For cosine-type fields, $R\left(\frac{-}{x}\right)$ has the MRD Property because the conditions are simply a restatement of (*).
- For mixed-type fields, $R\left(\frac{-}{x}\right)$ has the property because (**) holds. The $\cos(\phi_R)^2$ -component of (4.1b) is increasing by (c) and the $\sin(\phi_R)^2$ -component is strictly increasing because $e^{+cy}F(a-y) - e^{-cy}F(a+y)$ is strictly increasing by (b).
- For sine-type fields, the preceeding observation applied to (4.2a,b) insures that, for fixed λ_p , $N(\alpha_p, \lambda_p)$ is strictly decreasing as $|\alpha_p - \alpha_R|$ increases.

Theorem A.2 shows that sine-type elementary receptive field functions possessing a ZRD form a much broader class than the corresponding cosine-type and mixed-type field functions described by Theorem A.1. Similarly, sine-type elementary receptive field functions with the MRD Property form a broader class than the cosine-type and mixed-type field functions described by Theorem B.1. The following theorem gives partial results for sine-type fields.

Theorem B.3

Let the receptive field function $R\left(\frac{-}{x}\right)$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (3.5)). Assume $\cos(\phi_R) = 0$.

1. If $H_q(\rho) = e^{-c\rho^2}F(\rho^2)$ with $c > 0$ and $e^{-cy}F(y)$ is positive and strictly decreasing for all y , then $R\left(\frac{-}{x}\right)$ with $b_R = \cos(\phi_R) e^{-c}F(1)$ has the MRD Property. In this case, there is a unique ZRD given by $\alpha_{ZR} = \alpha_{MR} + \frac{\pi}{2}$.
2. If $H_q(\rho)$ is strictly decreasing on $0 < \rho < \sin(\zeta_R)$ and $H_q(\rho) = 0$ on $\sin(\zeta_R) < \rho < +\infty$, then $R\left(\frac{-}{x}\right)$ with $b_R = 0$ has the MRD Property. In this case, there is a sector of ZRDs and the sector is given by $0 < \zeta_R \leq |\alpha_{ZR} - \alpha_{MR}| \leq \frac{\pi}{2}$.
3. $R\left(\frac{-}{x}\right)$ has the MRD Property and more than one ZRD if and only if $H_q(\rho)$ satisfies (2) with $0 < \sin(\zeta_R) < 1$. In this case, $b_R = 0$ and there is a sector of ZRDs.

Result (1) follows by a direct check that, for fixed λ_p ,

$$N(\alpha_p, \lambda_p) = \left| H_q \left(\sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) - H_q \left(\sqrt{1 + 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) \right| \quad (4.4)$$

is strictly decreasing as $|\alpha_p - \alpha_R|$ increases. Result (2) follows by the same check after noting (4.4) reduces to a single term since one argument is necessarily 1. Result (3) follows from Theorem A.2, which characterizes the condition under which a sector of ZRDs can occur.

5a. Balanced Gabor Receptive Field Functions: Discussion and Examples

This section defines and discusses the balanced Gabor class of weights and elementary receptive field functions mentioned in Section 4. Two cases are analyzed: the simple

balanced Gabor receptive field function, corresponding to the traditional Gabor filter model of the receptive field, and a more general class of balanced Gabor weights, corresponding to elementary receptive fields with weight functions that are not simple gaussians, but that are products of gaussians and oscillatory components.

Balanced Gabor weights and receptive fields

Balanced Gabor weights (with exponent $\gamma_R > 0$) are weight functions $g(r)$ satisfying

$$g(r) \text{ is } C(0, +\infty), g(r) \text{ is absolutely integrable on } (0, +\infty), \text{ and } \int_0^\infty g(r) r dr = 1 \quad (5a.1a)$$

in which case the Hankel transform $H_g(\rho)$ satisfies

$$H_g(\rho) \text{ is } C[0, +\infty), H_g(0) = 1, \text{ and } H_g(+\infty) = 0 \quad (5a.1b)$$

and such that the following conditions hold:

- a. $H_g(\rho) = e^{-\gamma_R \rho^2} G(\rho^2)$ where $G(y)$ is continuous and positive on the real line, $G(0) = 1$, $G(y+1) = G(y)$;
- b. $e^{-\gamma_R y} G(y)$ is strictly decreasing on the real line;
- c. for each real a , $T_G(\gamma_R, a; y) := e^{-\gamma_R y} G(a-y) + e^{-\gamma_R y} G(a+y) - 2G(a)$ is initially zero on $0 \leq y \leq \delta_a$ and strictly increasing on $\delta_a \leq y < +\infty$;
- d. in (c), $\delta_a > 0$ if and only if the value a satisfies $e^{-cy} F(y) = e^{-ca} F(a) - K(a)(y-a)$ on $|y-a| \leq \delta_a$ for some constant $K(a) > 0$.

All parameters and independent variables are dimensionless. For each balanced Gabor Hankel transform (and thus for the corresponding weight function), the exponent is uniquely determined. The exponents thus partition or classify both the transforms and weights. For balanced Gabor weights $g(r)$, $g(0)$ is necessarily positive because (3.4) reduces to a positive integrand for $r=0$. *Simple balanced Gabor weights* refer to the special case $G(y) \equiv 1$. The following subsection discusses these weights and compares them with the traditional Gabor filter model for receptive fields.

Some points should be mentioned with regard to checking conditions (b), (c), and (d) of the definition. In checking (b) for a candidate periodic function $G(y)$, it is only necessary to check that $e^{-\gamma_R y} G(y)$ is strictly decreasing over an interval $[y_0, y_0 + 1]$ of length one, since shifting this function by a period multiplies it by a constant:

$e^{-\gamma_R(y+1)} G(y+1) = e^{-\gamma_R} e^{-\gamma_R y} G(y)$. In checking (c) for a candidate function $G(y)$, it is only necessary to check

1. the condition on $T_G(\gamma_R, a; y)$ for some interval $a_0 \leq a < a_0 + 1$ since the expression is 1-periodic in a ;
2. for given a , that $T_G(\gamma_R, a; y)$ is strictly increasing on $0 \leq y < 1$, since strictly increasing behavior on any unit interval implies a continued strict increase due to (b), which implies $e^{-\gamma_R y} G(-y)$ is strictly increasing, and the relation $T_G(\gamma_R, a; y+1) = e^{-\gamma_R} T_G(\gamma_R, a; y) + (e^{-\gamma_R} - e^{-\gamma_R}) e^{-\gamma_R y} G(-y)$.

Condition (d) says that the strictly decreasing function $e^{-cy} F(y)$ reduces to a linear function on some intervals. This condition (with $\delta_a > 0$) cannot hold for all a , or even for $0 \leq a \leq 1$, because $e^{-cy} F(y)$ would reduce to a strictly linear function over an interval larger than a period. Thus, $T_G(\gamma_R, a; y)$ must be strictly increasing on $0 \leq y < +\infty$ for some values of a .

While the sets of balanced Gabor Hankel transforms for different exponents are disjoint, the corresponding sets of periodic components are related:

Closure Lemma

Let

$$\Gamma(\gamma_R) := \{G(y) : G(y) \text{ satisfies conditions (a), (b), (c) for balanced Gabor weights with exponent } \gamma_R\}$$

Then:

1. $\Gamma(\gamma_R)$ is closed under convex combinations.
2. If $\gamma_1 < \gamma_2$, then $\Gamma(\gamma_1) \subsetneq \Gamma(\gamma_2)$.
3. $\bigcup_{0 < \gamma < \infty} \Gamma(\gamma)$ is closed under convex combinations.

Result (1) follows by a straightforward check of the conditions of the balanced Gabor definition when applied to convex combinations. Result (2) requires some manipulation and the proof is given in Appendix C. Result (3) follows directly from (1) and (2).

Two aspects of the Closure Lemma should be noted. First, the sets of periodic components for different exponents γ_R simply increase in extent as γ_R increases. Second, new weights can be formed by appropriately interpreted convex combinations of old weights. For example, let $g_1(r)$, $g_2(r)$ be balanced Gabor weights with corresponding Hankel transforms $e^{-\gamma_1(\rho^2)} G_1(\rho^2)$, $e^{-\gamma_2(\rho^2)} G_2(\rho^2)$, where $\gamma_1 < \gamma_2$. Then, by result (2), convex combinations of the form $H_3(\rho) = e^{-\gamma_2 \rho^2} (\alpha_1 G_1(\rho^2) + \alpha_2 G_2(\rho^2))$ with $\alpha_1, \alpha_2 \geq 0$ and $\alpha_1 + \alpha_2 = 1$ are Hankel transforms of new balanced Gabor weight functions $g_3(r)$.

A *balanced Gabor receptive field function* $R_{BG}(\vec{x})$ (with exponent γ_R) is an elementary receptive field function where the receptive field weight is a balanced Gabor function, that is,

$$R_{BG}(\vec{x}) := \frac{2\pi}{\lambda_R^2} g\left(2\pi \frac{|\vec{x}|}{\lambda_R}\right) \left(\cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) - b_R \right) \quad (5a.2)$$

where $g(r)$ is a balanced Gabor weight with associated exponent γ_R and $b_R = \cos(\phi_R) H_g(1) = \cos(\phi_R) e^{-\gamma_R}$. Notice this formula for the balance parameter b_R holds for *all* balanced Gabor fields. Here $\lambda_R > 0$ is the *spatial wavelength*, α_R is the *field orientation*,

and ϕ_R is the *phase shift*. The function $R_{BG}(\vec{x})$ is an elementary receptive field function with the MRD Property (and hence with the ZRD and Balanced Field Properties) for all values of the parameters, in particular, *all* values of the phase shift ϕ_R (Theorem B.2). Within the classes of Type I weights discussed in Section 4, balanced Gabor weights are precisely the weights that are phase-independent in the sense that elementary receptive fields based on these weights have the MRD Property for all field phase values ϕ_R .

Balanced Gabor receptive field functions have a unique ZRD given by $\zeta_R = \alpha_R + \frac{\pi}{2}$ (Theorem B.1).

The Fourier transform of a balanced Gabor receptive field function $R_{BG}(\vec{x})$, where the weight $g(r)$ has Hankel transform $H_g(\rho) = \mathbb{E}^{-\gamma_R \rho^2} G(\rho^2)$, follows from (3.6) and is given by

$$F_R(s_1, s_2) = \frac{1}{2} \exp\left(-\gamma_R \left(\lambda_R^2 |\vec{s}|^2 + 1\right)\right) \cdot \left[\mathbb{E}^{+i\phi_R} \exp\left(-2\gamma_R \lambda_R \vec{d}(\alpha_R) \cdot \vec{s}\right) G\left(\lambda_R^2 |\vec{s}|^2 + 2\lambda_R \vec{d}(\alpha_R) \cdot \vec{s}\right) + \mathbb{E}^{-i\phi_R} \exp\left(+2\gamma_R \lambda_R \vec{d}(\alpha_R) \cdot \vec{s}\right) G\left(\lambda_R^2 |\vec{s}|^2 - 2\lambda_R \vec{d}(\alpha_R) \cdot \vec{s}\right) - 2\cos(\phi_R) G\left(\lambda_R^2 |\vec{s}|^2\right) \right] \quad (5a.3)$$

The response $R \circ p$ of a balanced Gabor receptive field function to a sinusoidal grating pattern p , where the weight $g(r)$ has Hankel transform $H_g(\rho) = \mathbb{E}^{-\gamma_R \rho^2} G(\rho^2)$, follows from (3.8) and is given by

$$R \circ p = \frac{c_p}{2} \exp\left(-\gamma_R \left(1 + \frac{\lambda_R^2}{\lambda_p^2}\right)\right) \cdot \left(\cos(\phi_p - \phi_R) \left(\exp\left(+2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(\frac{\lambda_R^2}{\lambda_p^2} - 2\cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) - G\left(\frac{\lambda_R^2}{\lambda_p^2}\right) \right) + \cos(\phi_p + \phi_R) \left(\exp\left(-2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(\frac{\lambda_R^2}{\lambda_p^2} + 2\cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) - G\left(\frac{\lambda_R^2}{\lambda_p^2}\right) \right) \right) \quad (5a.4)$$

Notice the response $R \circ p$ does not reduce to a simple multiple of $\cos(\phi_p - \phi_R)$, in contrast to the bandlimited case (Section 5b), although this term is the exponentially dominant component. Consequently, for fixed grating orientation α_p , grating wavelength λ_p , and grating contrast c_p , the maximum of the response $R \circ p$ does not occur at a grating phase ϕ_p that exactly lines up with the field phase ϕ_R .

The max-response M of a balanced Gabor receptive field function, where the weight $g(r)$ has Hankel transform $H_g(\rho) = \mathbb{E}^{-\gamma_R \rho^2} G(\rho^2)$, to a sinusoidal grating pattern p follows from (3.9a,b) and is given by

$$M(c_p, \alpha_p, \lambda_p) = \max_{\phi_p} (R \circ p) = \frac{1}{2} |c_p| N(\alpha_p, \lambda_p) \quad (5a.5a)$$

with the amplitude factor $N(\alpha_p, \lambda_p) > 0$ given by

$$N(\alpha_p, \lambda_p)^2 = \exp\left(-2\gamma_R \left(\frac{\lambda_R^2}{\lambda_p^2} + 1\right)\right) \cdot \left[\cos(\phi_R)^2 N_c^2 + \sin(\phi_R)^2 N_s^2 \right] \quad (5a.5b)$$

where

$$N_c = \exp\left(+2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(-2\cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) + \exp\left(-2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(+2\cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) - 2G\left(\frac{\lambda_R^2}{\lambda_p^2}\right)$$

$$N_s = \exp\left(+2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(-2\cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) - \exp\left(-2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(+2\cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right)$$

Notice the max-response depends on the receptive field phase ϕ_R , in contrast to the bandlimited case (Section 5b). The maximum of the max-response, that is, the max of $N(\alpha_p, \lambda_p)$, occurs at $\alpha_p = \alpha_R$ for fixed λ_p , that is, the optimal grating orientation lines up with the receptive field orientation α_R . This is a general consequence of the MRD Property, but it can also be seen in the expression for $N(\alpha_p, \lambda_p)^2$, since the $\sin(\phi_R)^2$ -portion is strictly

decreasing as $|\alpha_p - \alpha_R|$ increases by condition (b) of the definition of balanced Gabor weight and the $\cos(\phi_R)^2$ -portion is strictly decreasing by condition (c). (Cosine-type filters may decrease to intervals where $N=0$ due to condition (d).) However, examples show that the maximum, for fixed grating orientation α_p , does not necessarily occur at $\lambda_p = \lambda_R$, that is, the optimal grating spatial wavelength does not necessarily occur at the receptive field spatial wavelength, another difference between the balanced Gabor and bandlimited cases (see examples below and Section 5b). Notice $N_s^2 - N_c^2 = 4 \left(-\exp\left(-2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(+2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) + G\left(\frac{\lambda_R^2}{\lambda_p^2}\right) \right) \left(\exp\left(+2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(-2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) + G\left(\frac{\lambda_R^2}{\lambda_p^2}\right) \right)$ because

$$N_s^2 - N_c^2 = 4 \left(-\exp\left(-2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(+2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) + G\left(\frac{\lambda_R^2}{\lambda_p^2}\right) \right) \left(\exp\left(+2\gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) G\left(-2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) + G\left(\frac{\lambda_R^2}{\lambda_p^2}\right) \right)$$

and each factor is positive because $\exp(-\gamma_R z) G(z+a)$ and $\exp(+\gamma_R z) G(-z+a)$ are strictly decreasing and strictly increasing functions of z by part (b) of the definition. In particular, the response of a pure sine-type receptive field will always be larger than the response of a pure cosine-type with an intermediate response for a mixed-type receptive field.

An explicit transform pair for the Hankel transform is

$$z(r) = \frac{1}{2} \frac{c+2\pi ik}{c^2+(2\pi k)^2} \exp\left(-\frac{r^2}{4} \frac{c+2\pi ik}{c^2+(2\pi k)^2}\right) \quad \text{and} \quad H_z(\rho) = e^{-c\rho^2+2\pi ik\rho^2} \quad (5a.6a)$$

Consequently, if the Hankel transform $H_q(\rho) = e^{-c\rho^2} G(\rho^2)$ has its periodic factor $G(y)$ expanded as a Fourier series, there is a corresponding explicit inversion:

$$q(r) = \sum_{k=-\infty}^{+\infty} \frac{g_k}{2} \frac{c+2\pi ik}{c^2+(2\pi k)^2} \exp\left(-\frac{r^2}{4} \frac{c+2\pi ik}{c^2+(2\pi k)^2}\right) \quad \text{for} \quad H_q(\rho) = e^{-c\rho^2} \sum_{k=-\infty}^{+\infty} g_k e^{+2\pi ik\rho^2} \quad (5a.6b)$$

The most basic instance of such explicit solutions, the case of a single harmonic for $G(y)$, is analyzed in a subsection below.

Example: Simple balanced Gabor receptive field functions (and traditional Gabor filters)

The traditional *Gabor filter* model for a receptive field function is

$$R\left(\frac{\vec{x}}{\sigma_R}\right) = \frac{1}{2\pi\sigma_R^2} = \frac{1}{2\pi\sigma_R^2} \exp\left(-\frac{|\vec{x}|^2}{2\sigma_R^2}\right) \cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) \quad (5a.7a)$$

where $\sigma_R > 0$ is the *spatial constant* and the other parameters correspond to (3.1). This is a typical normalization of the Gabor filter, where the gaussian factor has unit volume. Its Fourier transform is

$$F_R\left(\frac{\vec{s}}{\sigma_R}\right) = \exp\left(-2\pi^2\sigma_R^2\left(\left|\frac{\vec{s}}{\sigma_R}\right|^2 + \frac{1}{\lambda_R^2}\right)\right) \left[\cos(\phi_R) \cosh\left(4\pi^2\sigma_R^2 \frac{\vec{d}(\alpha_R) \cdot \vec{s}}{\lambda_R}\right) - i \sin(\phi_R) \sinh\left(4\pi^2\sigma_R^2 \frac{\vec{d}(\alpha_R) \cdot \vec{s}}{\lambda_R}\right) \right] \quad (5a.7b)$$

The traditional Gabor filter is not in general balanced. The exact value for the integral (2.2) is

$$\int_{\mathbb{R} \times \mathbb{R}} R(\vec{x}) d\vec{x} = \cos(\phi_R) \exp\left(-2\pi^2 \frac{\sigma_R^2}{\lambda_R^2}\right) \quad (5a.7c)$$

Consequently, it is balanced for sine-type filters ($\cos(\phi_R) = 0$) and is nearly zero when the dimensionless ratio λ_R / σ_R is sufficiently small. In the context of the Gabor filter, the ratio λ_R / σ_R is often interpreted as a measure of “bandwidth”. The value $\lambda_R / \sigma_R \approx 5/2$ is considered a typical mid-range value for simple cells (corresponding to a spatial frequency tuning bandwidth of approximately 1.7 octaves; DeValois & DeValois, 1988), giving 0.0425 for the exponential term in (5a.7c). That is, for this ratio, the imbalance of a cosine-type filter is about 4% of the gaussian component’s unit volume and becomes worse as λ_R / σ_R increases, that is, at large bandwidth.

The *simple balanced Gabor weight* $g(r)$ is determined by an exponent $\gamma_R > 0$. The weight and its Hankel transform $H_g(\rho)$ are given by

$$g(r) = \frac{1}{2\gamma_R} \exp\left(-\frac{r^2}{4\gamma_R}\right) \text{ with } H_g(\rho) = \exp(-\gamma_R \rho^2) \quad (5a.8)$$

The *simple balanced Gabor receptive field function* $R_{\text{SBG}}(\vec{x})$ and its Fourier transform are special cases of (5a.2,3):

$$R_{\text{SBG}}(\vec{x}) = \frac{2\pi}{\lambda_R^2} g\left(2\pi \frac{|\vec{x}|}{\lambda_R}\right) \left(\cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) - \cos(\phi_R) e^{-\gamma_R} \right) \quad (5a.9)$$

$$F_R(\vec{s}) = \exp\left(-\gamma_R \lambda_R^2 \left(|\vec{s}|^2 + \frac{1}{\lambda_R^2}\right)\right) \left[\cos(\phi_R) \left(\cosh\left(2\gamma_R \lambda_R \vec{d}(\alpha_R) \cdot \vec{s}\right) - 1 \right) - i \sin(\phi_R) \sinh\left(2\gamma_R \lambda_R \vec{d}(\alpha_R) \cdot \vec{s}\right) \right] \quad (5a.10)$$

Comparison of (5a.7) and (5a.9) shows that the receptive field function $R_{\text{SBG}}(\vec{x})$ is essentially the Gabor filter corrected to achieve balance. This is, of course, the motivation for the name of this class of weights. The exponent γ_R can be considered a dimensionless “shape” parameter which arises naturally in the derivation of these functions and unifies all cases of balanced Gabor weights, or, since matching the gaussian components of (5a.7) and (5a.9) gives the relation $\gamma_R = 2\pi^2(\sigma_R / \lambda_R)^2$, γ_R can be considered a measure of “reciprocal bandwidth”. Figure 1 illustrates weights $g(r)$ and corresponding Hankel transforms $H_g(\rho)$ for a range of values of γ_R . Figure 2 compares the traditional Gabor filter and the corresponding simple balanced Gabor receptive field function with identical gaussian components, that is, with identical weight functions (“envelopes”). Figure 2 also shows the corresponding Fourier transforms. The imbalance of the traditional Gabor field is shown by the nonzero values of its Fourier transform at the origin.

It will be useful to express our results in terms of the exponent γ_R since it provides a common unifying parameter for the entire class of balanced Gabor receptive field functions, both simple and nonsimple (see the following subsection). The typical value $\lambda_R / \sigma_R \approx 2.5$ gives $\gamma_R \approx 3.2$. Consequently, typical trial values for our plots in this subsection will be $\gamma_R = 6.0, 3.0, 1.5$ (respectively, $\lambda_R / \sigma_R = 1.8, 2.6, 3.6$) with the value $\gamma_R = 0.75$ ($\lambda_R / \sigma_R = 5.1$) included to indicate trends at more extreme values. Figure 2 shows that the two types of receptive fields essentially merge for large γ_R ($\gamma_R > 3$) and increasingly diverge for small γ_R ($\gamma_R < 3$) with $\gamma_R \approx 3$ as a nominal boundary between the two regimes. The same values

will be used in the following subsection for nonsimple balanced Gabor receptive field functions for comparison.

The response $R \circ p$ of a simple balanced Gabor receptive field function to a sinusoidal grating pattern p is a special case of (5a.4) and is given by

$$R \circ p = \frac{c_p}{2} \exp\left(-\gamma_R \left(1 + \frac{\lambda_R^2}{\lambda_p^2}\right)\right) \cdot \left(\cos(\phi_p - \phi_R) \left(\exp\left(+2 \gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) - 1 \right) + \cos(\phi_p + \phi_R) \left(\exp\left(-2 \gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) - 1 \right) \right) \quad (5a.11)$$

The behavior of this special case is the same as noted for nonsimple balanced Gabor behavior: the $\cos(\phi_p - \phi_R)$ -component is exponentially dominant at $\phi_p = \phi_R$, that is, where the grating phase matches the receptive field phase, but the maximum does not occur at exactly this value due to the $\cos(\phi_p + \phi_R)$ -component.

Similarly, the max-response function amplitude factor $N(\alpha_p, \lambda_p) > 0$ is a special case of (5a.5b) and is given by

$$N(\alpha_p, \lambda_p)^2 = 4 \exp\left(-2 \gamma_R \left(\frac{\lambda_R^2}{\lambda_p^2} + 1\right)\right) \cdot \left[\cos(\phi_R)^2 \left(\cosh\left(2 \gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) - 1 \right)^2 + \sin(\phi_R)^2 \left(\sinh\left(2 \gamma_R \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) \right)^2 \right] \quad (5a.12)$$

For fixed grating orientation α_p , the maximum for $N(\alpha_p, \lambda_p)$ does *not* necessarily occur at $\lambda_p = \lambda_R$, that is, where the grating wavelength matches the receptive field wavelength. This behavior is illustrated in Figure 3. Plots of $N(\alpha_p, \lambda_p)$ vs. λ_p are given for the optimal orientation $\alpha_p = \alpha_R$. These plots are essentially the same for all receptive field phases at $\gamma_R = 6$ and with maxima essentially at $\lambda_p = \lambda_R$, but show increasing dependence on the receptive field phase ϕ_R and increasing deviation of the optimal grating wavelength from the receptive field wavelength λ_R as γ_R decreases. This behavior is quite different from the bandlimited case, where the max-response is completely independent of ϕ_R (Section 5b).

For fixed grating wavelength λ_p , $N(\alpha_p, \lambda_p)$ has a maximum at $\alpha_p = \alpha_R$, that is, where the grating orientation matches the field orientation, due to the MRD Property. Figure 3 also shows such plots of response vs. grating orientation at optimal grating wavelengths. The increasing dependence of the response on the receptive field phase ϕ_R as γ_R decreases is again evident.

Example: A class of nonsimple balanced Gabor receptive field functions

We define *general balanced Gabor weights of order 1* (with exponent γ_R) as those weights given by

$$g(r) = \frac{1}{2 \gamma_R (1 + c_R \cos(\psi_R))} \exp\left(-\frac{r^2}{4 \gamma_R}\right) + \frac{c_R \gamma_R}{2(1 + c_R \cos(\psi_R))(\gamma_R^2 + 4 \pi^2)} \exp\left(\frac{-\gamma_R r^2}{4(\gamma_R^2 + 4 \pi^2)}\right) \left(\cos\left(\frac{\pi r^2}{2(\gamma_R^2 + 4 \pi^2)} + \psi_R\right) + \frac{2 \pi}{\gamma_R} \sin\left(\frac{\pi r^2}{2(\gamma_R^2 + 4 \pi^2)} + \psi_R\right) \right) \quad (5a.13a)$$

with Hankel transform given by

$$H_g(\rho) = e^{-\gamma_R \rho^2} G(\rho^2) \quad \text{where} \quad G(y) = \frac{1 + c_R \cos(2 \pi y - \psi_R)}{1 + c_R \cos(\psi_R)} \quad (5a.13b)$$

where the coefficient c_R satisfies the constraint

$$|c_R| \leq \frac{\gamma_R^2}{\gamma_R^2 + 4\pi^2} \quad (5a.13c)$$

In the Hankel transform, notice that $G(y)$ is the most general form of periodic function with a single harmonic term that satisfies the definition for a balanced Gabor weight, and it can be shown that the bound on $|c_R|$ is a necessary and sufficient condition to satisfy conditions (a), (b), (c) of that definition (Appendix C). Condition (d) is vacuous for this case. The simple balanced Gabor occurs for $c_R = 0$ and can be considered the “order 0” form. With the Hankel transform $H_g(\rho)$ determined, the inverse transform $g(r)$ follows by the explicit inversion formula (5a.6a).

In $g(r)$, notice the oscillatory component is the more slowly decaying component:

$$g(r) = \frac{1}{2\gamma_R(1+c_R\cos(\psi_R))} \exp\left(\frac{\gamma_R r^2}{4(\gamma_R^2 + 4\pi^2)}\right) \cdot \left[\exp\left(-\frac{\pi^2 r^2}{\gamma_R(\gamma_R^2 + 4\pi^2)}\right) + \frac{c_R \gamma_R^2}{\gamma_R^2 + 4\pi^2} \left(\cos\left(\frac{\pi r^2}{2(\gamma_R^2 + 4\pi^2)} + \psi_R\right) + \frac{2\pi}{\gamma_R} \sin\left(\frac{\pi r^2}{2(\gamma_R^2 + 4\pi^2)} + \psi_R\right) \right) \right] \quad (5a.13d)$$

The *balanced Gabor receptive field function* (order 1) $R_{\text{SBG}}(\vec{x})$ is the special case of (5a.2) given by

$$R_{\text{BG}}(\vec{x}) = \frac{2\pi}{\lambda_R^2} g\left(2\pi \frac{|\vec{x}|}{\lambda_R}\right) \left(\cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) - \cos(\phi_R) e^{-\gamma_R} \right) \quad (5a.14)$$

where $g(r)$ is given by (5a.13a). In the same way, the Fourier transform $F_R(s_1, s_2)$, the response function RO_p for sinusoidal grating patterns p , and the max-response amplitude factor $N(\alpha_p, \lambda_p) > 0$ are given by formulas combining (5a.13) with the general formulas (5a.3), (5a.4), and (5a.5b), respectively.

The figures show examples of these nonsimple balanced Gabor receptive field functions and their behavior. There are two sets: Figures 4, 5, 6 have $\psi_R = 0$ (rad) and 7, 8, 9 have $\psi_R = \pi/2$ (rad). Each plot shows behavior for exponents $\gamma_R = 0.75, 1.5, 3.0, 6.0$ and for cases of $G(y)$ with the coefficient c_R taking the value 0 (that is, the simple balanced Gabor case) and the two most extreme values permitted by the bound (5a.13c). Since this bound is small for small γ_R , the plots for $\gamma_R = 0.75$ essentially reduce to the simple balanced Gabor case. As γ_R increases, however, the bound also increases, and substantial differences from the simple balanced Gabor are evident at $\gamma_R = 6.0$. Specifically, Figures 4 and 7 show the Hankel transforms $H_g(\rho)$ and weight functions $g(r)$ for these parameter values. Notice that the weight functions $g(r)$ can become negative, in contrast to the simple balanced Gabor where the weights are strictly positive. Figures 5 and 8 show the corresponding receptive field functions (with the weight functions as envelopes) and Fourier transforms for the receptive field functions (cosine-type fields). Figures 6 and 9 show plots of $N(\alpha_p, \lambda_p)$ vs. λ_p at the optimal grating orientation ($\alpha_p = \alpha_R$). The optimal grating wavelength is quite different from the receptive field wavelength λ_R for small γ_R but begins to match it as γ_R increases. At the same time, the variation in these curves increases with γ_R . That is, as bandwidth decreases and the simple balanced Gabor and traditional Gabor receptive fields become identical (see Figure 2), the nonsimple balanced Gabor weights and fields show a widening range of behavior. The same figures show plots of $N(\alpha_p, \lambda_p)$ vs. α_p at optimal wavelengths λ_p . The maximum response occurs at $\alpha_p = \alpha_R$, that is, when the grating orientation matches

the receptive field orientation, in agreement with the Maximum Response Direction Property, and the variation in the response increases with γ_R .

5b. Bandlimited Field Functions: Discussion and Examples

This section defines and discusses the bandlimited class of weights and elementary receptive field functions mentioned in Section 4. A bandlimited class with explicit analytic solutions (Bessel fields) is given as an example.

Bandlimited weights and receptive fields

Bandlimited weights (with support $0 < s_R \leq 1$) are weight functions $b(r)$ satisfying

$$b(r) \text{ is } C(0, +\infty), \quad b(r)r \text{ is absolutely integrable on } (0, +\infty), \quad \text{and} \quad \int_0^\infty b(r)r dr = 1 \quad (5b.1a)$$

in which case the Hankel transform $H_b(\rho)$ satisfies

$$H_b(\rho) \text{ is } C[0, +\infty), \quad H_b(0) = 1, \quad \text{and} \quad H_b(+\infty) = 0 \quad (5b.1b)$$

and such that there exists a value $0 < s_R \leq 1$ with the properties

- a. $H_b(\rho) = 0$ on $s_R \leq \rho < +\infty$;
- b. $H_b(\rho)$ is strictly decreasing on $0 \leq \rho < s_R$ with $H_b(0) = 1$.

All parameters and independent variables are dimensionless. For each bandlimited Hankel transform (and thus for the corresponding weight function), the support value s_R is uniquely determined. The support values thus partition or classify both the transforms and the weights. Since the support s_R is a measure of the width of the Hankel transform and thus of the Fourier transform for the receptive field (see below), it corresponds to a measure of the bandwidth of the bandlimited receptive field function. For bandlimited weights $b(r)$, $b(0)$ is necessarily positive (because (3.4) reduces to a positive integrand for $r = 0$). In contrast to balanced Gabor weights, no explicit analytic structure occurs naturally for bandlimited weights, and no analog for “simple” balanced Gabor weights occurs.

A closure result can be stated for the bandlimited case and takes a simpler form than for the balanced Gabor case:

Closure Lemma

1. The set of bandlimited Hankel transforms and the corresponding set of bandlimited weights for a fixed support value s_R are closed under convex combinations.
2. The set of all bandlimited Hankel transforms and the corresponding set of all bandlimited weights are closed under convex combinations.

Both parts follow by a straightforward check of the conditions of the bandlimited definition when applied to convex combinations.

If $u(r)$ is a bandlimited weight function with support $s_R = 1$, it follows directly from the definition that, for an arbitrary c , $0 < c \leq 1$, the function $b(r) = c^2 u(cr)$ is a bandlimited

weight with support $s_R = c$ and Hankel transform $H_b(\rho) = H_u\left(\frac{\rho}{c}\right)$. That is, scaling maps bandlimited weights with a given support value to weights with other support values. In particular, the equivalence class of weights for a given support value s_R is a scaled copy of any other equivalence class.

A *bandlimited receptive field function* $R_{BL}(\vec{x})$ is an elementary receptive field function where the field weight is a bandlimited function, that is, $q(r)$ has been replaced by $b(r)$,

$$R_{BL}(\vec{x}) = \frac{2\pi}{\lambda_R^2} b\left(2\pi \frac{|\vec{x}|}{\lambda_R}\right) \cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) \quad (5b.2)$$

The balance parameter vanishes because $b_R = \cos(\phi_R) H_g(1) = 0$ for *all* bandlimited fields. Bandlimited receptive field functions have the MRD Property (and hence with the ZRD and Balanced Field Properties) because bandlimited weights are identical with the Type II class of weights for elementary receptive field functions defined by Theorem B.1.

The support value s_R determines the corresponding sector of ZRDs (possibly reducing to a single line) by $s_R = \sin(\zeta_R)$. The sector is centered on $\alpha_R + \frac{\pi}{2}$ and is bounded by ζ_R , that is, the ZRD sector consists of $\zeta_R \quad \alpha_P - \alpha_R \quad \pi - \zeta_R$.

The Fourier transform of a bandlimited receptive field function reduces to

$$F_R(s_1, s_2) = \frac{1}{2} \left(e^{+i\phi_R} H_b \left(\sqrt{(\lambda_R s_1 + \cos(\alpha_R))^2 + (\lambda_R s_2 + \sin(\alpha_R))^2} \right) + e^{-i\phi_R} H_b \left(\sqrt{(\lambda_R s_1 - \cos(\alpha_R))^2 + (\lambda_R s_2 - \sin(\alpha_R))^2} \right) \right) \quad (5b.3)$$

For any point (s_1, s_2) , at least one of the arguments is ≤ 1 , so the sum always reduces to a single term. The motivation for the name *bandlimited* comes from the fact that the Fourier transforms of the weight and receptive field functions have compact support.

The response $R \circ p$ of a bandlimited receptive field function to a sinusoidal grating pattern p is given by (assuming $|\alpha_P - \alpha_R| \leq \frac{\pi}{2}$)

$$R \circ p = \frac{c_P}{2} \cos(\phi_P - \phi_R) H_b \left(\sqrt{1 - 2 \cos(\alpha_P - \alpha_R) \frac{\lambda_R}{\lambda_P} + \frac{\lambda_R^2}{\lambda_P^2}} \right) \quad (5b.4)$$

In contrast to the balanced Gabor case, the response $R \circ p$ reduces to a simple multiple of $\cos(\phi_P - \phi_R)$. Thus, for fixed grating orientation α_P , grating wavelength λ_P , and grating contrast c_P , the maximum of the response $R \circ p$ necessarily occurs when the grating phase ϕ_P matches the field phase ϕ_R .

The max-response of a bandlimited receptive field function to a sinusoidal grating pattern is given by

$$M(c_P, \alpha_P, \lambda_P) = \max_{\phi_P} (R \circ p) = \frac{1}{2} |c_P| N(\alpha_P, \lambda_P) \quad (5b.5a)$$

where the amplitude factor $N(\alpha_P, \lambda_P)$ for the max-response is given by

$$N(\alpha_P, \lambda_P) = H_b \left(\sqrt{1 - 2 \cos(\alpha_P - \alpha_R) \frac{\lambda_R}{\lambda_P} + \frac{\lambda_R^2}{\lambda_P^2}} \right) \quad (5b.5b)$$

The simplicity of the formula for $N(\alpha_P, \lambda_P)$ provides several results:

1. The maximum of the max-response amplitude factor is $N = 1$ and occurs for exactly one pair of values, $(\alpha_p, \lambda_p) = (\alpha_R, \lambda_R)$.
2. The response function is *independent* of the field phase ϕ_R .
3. If the grating direction α_p is fixed, then $|\alpha_p - \alpha_R| < \zeta_R$, where $s_R = \sin(\zeta_R)$, is required for a nonzero response. A nonzero response occurs only for wavelengths λ_p in the interval

$$\cos(\alpha_p - \alpha_R) - \sqrt{\cos(\alpha_p - \alpha_R)^2 - \cos(\zeta_R)^2} \leq \frac{\lambda_R}{\lambda_p} \leq \cos(\alpha_p - \alpha_R) + \sqrt{\cos(\alpha_p - \alpha_R)^2 - \cos(\zeta_R)^2}$$

which narrows to a single point when $\alpha_p - \alpha_R = \zeta_R$. The max over this interval

occurs at the midpoint $\frac{\lambda_R}{\lambda_p} = \cos(\alpha_p - \alpha_R)$ and that max-value in this fixed direction is given by $N = H_b(\sin(\alpha_p - \alpha_R))$.

4. If the grating wavelength λ_p is fixed, then $\left| \frac{\lambda_R}{\lambda_p} - 1 \right|$, where $s_R = \sin(\zeta_R)$, is required for a nonzero response. A nonzero response occurs only for orientations α_p in the range $\cos(\alpha_p - \alpha_R) \geq \frac{1}{2} \left(\frac{\lambda_R}{\lambda_p} + (\zeta_R)^2 \frac{\lambda_p}{\lambda_R} \right)$ and the maximum response within this interval is at the midpoint $\alpha_p = \alpha_R$ giving $N = H_b \left(\left| \frac{\lambda_R}{\lambda_p} - 1 \right| \right)$.

- a. The broadest interval of nonzero response occurs when $\frac{\lambda_R}{\lambda_p} = \cos(\zeta_R) < 1$, in which case $|\alpha_p - \alpha_R| < \zeta_R$ (that is, extends to the ZRD sector) with maximum value $H_b(1 - \cos(\zeta_R))$.

- b. The interval of nonzero response for the optimal wavelength $\frac{\lambda_R}{\lambda_p} = 1$ is $\left| \sin \left(\frac{\alpha_p - \alpha_R}{2} \right) \right| \leq \frac{1}{2} \sin(\zeta_R)$ with maximum value $H_b(0) = 1$.

Notice the broadest interval is larger than the interval at the optimal wavelength.

The synthesis/analysis formulas for the Hankel transform and the fact that $H_b(\rho)$ vanishes outside the interval $[0, s_R]$ with $0 < s_R \leq 1$ for the bandlimited class, give a general formula for the weights:

$$b(r) = \int_0^1 J_0(\rho r) H_b(\rho) \rho d\rho = \sum_{k=0}^{\infty} \frac{(-r^2/4)^k}{(k!)^2} \int_0^{s_R} H_b(\rho) \rho^{2k+1} d\rho \quad (5b.6)$$

Example: Bessel fields

An example of the explicit inversion (5b.6) is the following Hankel transform and inverse transform. These functions satisfy the conditions for a bandlimited weight function with support parameter $s_R = 1$: for $\nu_R > 3/2$,

$$H_b(\nu_R; \rho) := \begin{cases} (1 - \rho^2)^{\nu_R - 1} & \text{for } 0 \leq \rho < 1 \\ 0 & \text{for } 1 \leq \rho \end{cases} \quad (5b.7a)$$

$$b(\nu_R; r) := 2^{\nu_R - 1} \Gamma(\nu_R) r^{-\nu_R} J_{\nu_R}(r) = \int_0^1 J_0(\rho r) (1 - \rho^2)^{\nu_R - 1} \rho d\rho \quad (5b.7b)$$

As ν_R increases, the transform H_b decreases more rapidly to zero and has a smoother transition at $\rho = 1$. The weight function $b(\nu_R; r)$ is positive at the origin with $b(\nu_R; 0) = (2\nu_R)^{-1}$, has infinitely many positive roots, and has a first root at $r = j_{\nu_R, 1} \sim \nu_R + 1.8557571 \nu_R^{1/3} + O(\nu_R^{-1/3})$ as $\nu_R \rightarrow \infty$ (Abramowitz & Stegun, 1972). The function satisfies $b(\nu_R; r) = O(r^{-\nu_R - 1/2})$ as $r \rightarrow \infty$, and the condition $\nu_R > 3/2$ insures the absolute integrability of $b(r)$ as well as continuity of the Hankel transform.

Scaling gives corresponding weight functions with arbitrary support $0 < s_R \leq 1$:

$$b(r) = s_R^{-2} b(\nu_R; s_R r) \quad \text{and} \quad H_b(\rho) = H_b\left(\nu_R; \frac{\rho}{s_R}\right) \quad (5b.7c)$$

This set of scaled functions $b(r)$ will be called *Bessel weights of order ν_R* , referring to the order of the corresponding Bessel function. The support parameter s_R determines an equivalence class of weights.

The *Bessel receptive field functions* $R_B(\vec{x})$ are the elementary receptive field functions given by

$$R_B(\vec{x}) = 2\pi \frac{s_R^2}{\lambda_R^2} b\left(\nu_R; 2\pi \frac{s_R}{\lambda_R} |\vec{x}|\right) \cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) \quad (5b.8)$$

where $0 < s_R \leq 1$ is the support parameter. Increasing the order ν_R broadens and flattens the weight factor b . As noted under the general discussion of bandlimited fields, these fields have the MRD Property (and hence with the ZRD and Balanced Field Properties). The boundary angle ζ_R between the sector of nonzero response and the sector of ZRDs is given by $s_R = \sin(\zeta_R)$.

Formulas and properties for the Fourier transform, response function, and max-response for Bessel fields carry over directly from the general discussion. We note a couple of simplifications. The response function $R \circ p$ for a sinusoidal grating pattern p reduces to

(assuming $|\alpha_p - \alpha_R| \leq \frac{\pi}{2}$)

$$\begin{aligned} R \circ p &= \frac{c_p}{2} \cos(\phi_p - \phi_R) H_b\left(\nu_R; \frac{1}{s_R} \sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}}\right) \\ &= \frac{c_p}{2} \frac{\cos(\phi_p - \phi_R)}{s_R^2 \nu_R^{-2}} \left[s_R^2 - \sin(\alpha_p - \alpha_R)^2 - \left(\frac{\lambda_R}{\lambda_p} - \cos(\alpha_p - \alpha_R) \right)^2 \right]^{\nu_R - 1} \\ &\quad \text{when } \sin(\alpha_p - \alpha_R)^2 + \left(\frac{\lambda_R}{\lambda_p} - \cos(\alpha_p - \alpha_R) \right)^2 < s_R^2 \\ &= 0 \text{ otherwise} \end{aligned} \quad (5b.9)$$

and the max-response to a sinusoidal grating pattern has amplitude factor $N(\alpha_p, \lambda_p)$ given explicitly by

$$\begin{aligned}
N(\alpha_p, \lambda_p) &= H_b \left(\nu_R, \frac{1}{s_R} \sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}} \right) \\
&= \frac{1}{s_R^{2\nu_R-2}} \left[s_R^2 - \sin(\alpha_p - \alpha_R)^2 - \left(\frac{\lambda_R}{\lambda_p} - \cos(\alpha_p - \alpha_R) \right)^2 \right]^{\nu_R-1} \quad (5b.10) \\
&\quad \text{when } \sin(\alpha_p - \alpha_R)^2 + \left(\frac{\lambda_R}{\lambda_p} - \cos(\alpha_p - \alpha_R) \right)^2 < s_R^2 \\
&= 0 \text{ otherwise}
\end{aligned}$$

Figures 10, 11, 12 show examples of Bessel receptive field functions and their behavior. Each plot shows curves for orders $\nu_R = 2.0, 3.5, 5.0$. Each figure shows plots for support values $s_R = 0.5, 0.7, 0.85, 1.0$, corresponding to scaled versions of each other. Figure 10 shows Hankel transforms and corresponding weight functions. Figure 11 shows Fourier transforms and corresponding receptive field functions. Since the balance parameter $b_R = 0$ for the bandlimited case, the receptive field curves always match up with their envelopes and the nodes always match exactly with the zeros of the periodic carrier. Figure 12 shows amplitude factors $N(\alpha_p, \lambda_p)$ at optimal orientations ($\alpha_p = \alpha_R$) and optimal wavelengths ($\lambda_p = \lambda_R$). Notice, in contrast to the balanced Gabor case, the max-response is *independent* of the receptive field phase ϕ_R .

6. Summary and General Discussion

First, we suggest three properties as theoretical postulates for simple cell receptive fields (Section 2). Experimental studies have established that simple cells show three well-defined properties, formulated here as the *Balanced Field Property* (spatially homogeneous patterns produce a zero response); the *Zero Response Direction Property* (there is a direction, i.e., stimulus orientation, which elicits a zero response to sinusoidal gratings); and the *Maximum Response Direction Property* (the maximum response to sinusoidal gratings decreases monotonically to zero as direction, i.e., stimulus orientation, changes from the optimal direction to a direction perpendicular to the optimal). These properties are directly motivated by well-established experimentally observed behavior (Hubel & Weisel, 1968; 1977; DeValois & DeValois, 1988; Ringach, 2002). Since our analysis is based on responses to sinusoidal grating stimuli, these properties can be restated in terms of the Fourier transform of the simple cell receptive field function. For the Fourier transform $F_R(s_1, s_2)$ of the receptive field function $R(x_1, x_2)$, the properties are:

1. *Balanced Field Property*: $F_R(0, 0) = 0$.
2. *Zero Response Direction Property*: There exists a direction α_{ZR} such that $F_R(s \cos(\alpha_{ZR}), s \sin(\alpha_{ZR})) = 0$ for all s .
3. *Maximum Response Direction Property*: There exists a direction α_{MR} such that, for each fixed s , either $F_R(s \cos(\alpha), s \sin(\alpha))$ is zero for all α or strictly decreases to zero (and remains zero) as $|\alpha - \alpha_{MR}|$ increases.

Second, we demonstrate that these properties can serve as a productive basis for theoretical development. A simple role for such postulates is to provide definite criteria that any proposed receptive field model should satisfy, where such models should either satisfy the criteria or, if not, should possess compensating advantage(s). For example, the traditional Gabor filter is not balanced, but the imbalance is negligible over a certain parameter range and the filter provides a useful explicit approximation to observed receptive field functions within that range. A deeper role is to provide a catalyst. We have demonstrated such a role by using the properties to completely characterize *elementary receptive field functions*, that is, to characterize the weight functions q defining such fields for cosine-type and mixed-type fields (Sections 3, 4). This derivation began by obtaining a description of receptive field functions with the ZRD Property (Appendix A), then refining that description to obtain

receptive field functions with the MRD Property (Appendix B). The usefulness of this stepwise approach motivates the above formulation of three cumulative properties.

This characterization yields two disjoint classes of elementary receptive field functions, the *balanced Gabor* class (Section 5a) and the *bandlimited* class (Section 5b). The balanced Gabor class appears to be an essentially new class of receptive field models, although its most basic case, the *simple balanced Gabor receptive field function*, is a straightforward modification of the traditional Gabor filter. Bandlimited models for receptive fields have been proposed (Victor & Knight, 2003), and the bandlimited class derived here is an unexpected connection with that literature. One difference between the classes is that the balanced Gabor class has a unique zero response direction (orthogonal to the maximum response direction) while the bandlimited class can have a sector of zero response directions/orientations (centered on an orientation orthogonal to the maximum response orientation, but possibly reducing to a single orientation). Another difference is that the Hankel transform $H_q(\rho)$ of the weight function q is strictly positive for the balanced Gabor class (although it decays exponentially fast due to a characteristic gaussian factor) while $H_q(\rho)$ for the bandlimited case is nonnegative and is always zero outside some bounded region (i.e., has compact support).

Third, we provide explicit examples and short studies of these two classes of elementary receptive field functions (Sections 5a, 5b). It should be noticed that the balanced Gabor class has a partially determined analytic structure (the Hankel transform $H_q(\rho)$ of the weight function is the product of a gaussian and a periodic function); while equally well-defined, the bandlimited class does not possess a comparable analytic form for the Hankel transform.

With regard to further work, the fact that a set of three experimentally based properties leads to two *disjoint* classes of elementary receptive field functions raises the obvious question of which class better describes the simple cell receptive field. Consequently, the most immediate question raised by this paper is whether experimental study might lead to a decision between these two classes. In addition, there are natural directions for generalizing both the results obtained here and the approach. We set these out under the following suggestions.

Experimental diagnostics for receptive field functions—The two classes of receptive field functions resulting from our analysis necessarily exhibit similar behavior since they both satisfy the three imposed experimentally-based response properties to sinusoidal grating stimuli. Not surprisingly, therefore, identifying diagnostic differences in the responses of the two classes of elementary receptive field functions that are both experimentally detectable and empirically conclusive may prove challenging. The construction of explicit examples of receptive field functions of both types, based on the selection of receptive field parameter values which result in physiologically plausible behavior, has suggested some differences which *may* have experimental significance. We list some of these differences by way of example but note that further study is required:

- a. Bandlimited receptive field functions can have a sector of Zero Response Directions, whereas balanced Gabor receptive fields always possess a unique Zero Response Direction. While a narrow range of zero response directions (orientations) may be experimentally indistinguishable from a single Zero Response Direction, a broad range of grating orientations yielding zero response points to bandlimited receptive field structure.
- b. Bandlimited receptive field functions have max-response (i.e., tuning) functions with respect to both spatial frequency and orientation that are completely independent of the receptive field phase shift ϕ_R . By contrast, balanced Gabor max-

response functions depend on receptive field phase, and max-response values for sine-type receptive field functions are always larger than those for cosine-type receptive field functions. In addition, initial numerical calculations suggest that, for balanced Gabor types, cosine phase receptive field functions are more narrowly tuned for both spatial frequency and orientation than are mixed or sine phase receptive fields (see Figs. 3, 6 & 9). Population level surveys of simple cell behavior might be diagnostic if, as a group, simple cell max-responses, orientation, or spatial frequency tuning were found to systematically vary with receptive field phase, since this would contraindicate the bandlimited class of receptive field model for which these response properties are invariant with receptive field phase.

- c. The weight function, q , of the simple balanced Gabor receptive field function is always non-negative, whereas the weight functions of nonsimple balanced Gabor and bandlimited types may assume negative values. When the weight function changes sign the phase of the periodic carrier function undergoes an abrupt reversal. Such abrupt carrier phase reversals may be of diagnostic value since discovering simple cells with such receptive field structure would disqualify the simple balanced Gabor receptive field type.

Theoretical diagnostics for evaluating receptive field functions—There may be theoretical grounds, such as minimal energy or other optimality arguments, for preferring one class of elementary receptive field function over the other. We have not extended our analysis to include such considerations, but this may be a fruitful direction for further study.

Further developments of the present analysis—The present analysis has assumed that the weight function, q , of the elementary receptive field function is circularly symmetric and satisfies the regularity conditions described in (3.2), in particular that $\int_0^\infty q(r) r dr = 1$ (viz., that the weight function possesses unit volume) which in turn implies that the Hankel transform satisfies $H_q(\rho) = 1$. This is a broad normalization condition since a simple scaling of the weight function can transform *any* nonzero weight integral to unit value. Several questions regarding further generalizations naturally arise:

- a. What receptive field structures might satisfy the Balanced Field, Zero Response Direction and Maximum Response Direction constraints if our assumed regularity condition $\int_0^\infty q(r) r dr = 1$ is modified to be the *singular condition* $\int_0^\infty q(r) r dr = 0$, in which case the weight function itself must integrate to zero? If valid receptive field functions result, that is, if nontrivial weight functions, q , are discovered which satisfy the constraints imposed by the three postulated response properties, then they would necessarily form a new class of receptive field structures distinct from the two classes described in this paper, since no scaling of independent or dependent variables can convert the singular to the regular condition.
- b. What are the implications for our analysis of receptive field functions if the condition of a circularly symmetric weight function, q , is relaxed to allow for *elliptically* symmetric weights? Both psychophysical (Foley, Varadharajan, Koh & Farias, 2007; Polat & Tyler, 1999; Anderson & Burr, 1991; but see Watson, Barlow & Robson, 1983) and physiological observations (Parker & Hawken, 1988; Ringach, 2002) indicate that simple cell receptive fields with elliptically symmetric weight functions may commonly occur. This suggests that extending the current analysis to include such receptive field variations would be profitable. This extension of our analysis is, however, beyond the scope of the present paper.
- c. We have defined elementary receptive field functions in (3.1) to be the product of a circularly symmetric weight and a simple periodic carrier. The simple periodic

carrier corresponds to the first two terms of a Fourier series expansion. Can receptive field functions with a general periodic carrier be characterized? That is, can non-elementary receptive field functions which satisfy our postulated response behaviors be characterized, where the functions have the following form, and $p(y)$ is an arbitrary 1-periodic function:

$$R(\vec{x}) := \frac{2\pi}{\lambda_r^2} q\left(2\pi \frac{|\vec{x}|}{\lambda_r}\right) p\left(\frac{\vec{d}(\alpha_r) \cdot \vec{x}}{\lambda_r}\right) \quad (6.1)$$

- d. Our approach has been to deduce the mathematical implications of three simple, but generally accepted, properties of the response of simple cells to sinusoidal gratings. Our analysis relies, in part, on the fact that statements about such responses convert directly to statements about the Fourier transform of elementary receptive field functions, and on the leverage afforded by the intersection of constraints imposed by these three response properties. It is therefore of interest to ask whether there might be additional properties of simple cells, well-defined but perhaps less widely recognized (and not necessarily restricted to sinusoidal grating stimuli) that could be incorporated to supplement the three postulated properties? It should be kept in mind that such properties might be considered of minor or secondary importance from the viewpoint of experimental significance, yet might have unexpected power in mathematical terms. For example, the Zero Response Direction property might easily be regarded as of secondary importance in comparison with the Maximum Response Direction property, yet studying the implications of the existence of a zero response direction provided important initial results for this work. The incorporation of such an additional response property to the present analysis would quite probably determine which class of elementary receptive field function best describes simple cell receptive fields.

Classical versus Extraclassical Receptive Fields—A final caveat is that our analysis is motivated by the response properties of the so-called classical receptive field of V1 neurons, which typically demonstrate the Balanced Field Property. A large body of literature suggests that V1 neurons also possess “extraclassical” receptive fields (Albright & Stoner, 2002), and a sizable proportion of V1 neurons respond to homogeneous luminance modulations if these stimuli extend into the extraclassical region, which may subtend many degrees of visual angle beyond the classical receptive field (Kayama, Riso, Bartlett & Doty, 1979; Kinoshita & Komatsu, 2001; MacEvoy, Kim & Paradiso, 1998; Peng & Van Essen, 2005; Roe, Lu & Hung, 2005; Rossi and Paradiso, 1996; 1999; Rossi, Rittenhouse & Paradiso, 1996; Wachtler, Sejnowski & Albright, 2003).

Acknowledgments

This work was supported by grants R01 EY014015 and NIH P20 RR020151. The National Center for Research Resources (NCRR) and the National Eye Institute (NEI) are components of the National Institutes of Health (NIH). The contents of this report are solely the responsibility of the authors and do not necessarily reflect the official views of the NIH, NCRR, or NEI.

Appendix A. Proofs for Lemma A and Theorems A.1 and A.2

This Appendix states and proves a key result, Lemma A, and applies it to prove Theorems A.1 and A.2 (Section 4).

It will be useful to extend the domain of $H_q(\rho)$ to the whole real line and rephrase Lemma 2. Since $q(r)$ satisfies (3.2), the Hankel transform $H_q(\rho)$, defined by (3.4), exists for all real ρ and is even because $J_0(x)$ is even. Properties (3.5) become

$$H_q(\rho) \text{ is } C(-\infty, +\infty), H_q(0)=1, \text{ and } H_q(\pm\infty)=0 \quad (\text{A.1})$$

With respect to Lemma 2, we can assume $a_{ZR} - a_R$ is restricted to $[0, +\pi/2]$ because conditions (3.11a,b) are unchanged when $a_{ZR} - a_R$ is translated by $\pm\pi$ and the cosine is even. Conditions (3.11a,b) are also unchanged when $x = \lambda_R / \lambda_P$ is replaced by $-x$ (and hold at $x = 0$ by continuity). Lemma 2 can therefore be restated as:

Lemma 2'

Let the receptive field function $R\left(\frac{-}{x}\right)$ be given by (3.1). Let $q(r)$ satisfy (3.2) (in which case the Hankel transform $H_q(\rho)$ exists and satisfies (A.1)). Then $R\left(\frac{-}{x}\right)$ has the ZRD Property if and only if there exists a value $\beta_R, 0 \leq \beta_R \leq \pi/2$, such that the following relations hold for $-\infty < x < +\infty$:

$$\cos(\phi_R) \left(H_q \left(\sqrt{1 - 2 \cos(\beta_R) x + x^2} \right) + H_q \left(\sqrt{1 + 2 \cos(\beta_R) x + x^2} \right) \right) = 2 b_R H_q(x) \quad (\text{A.2a})$$

$$\sin(\phi_R) \left(H_q \left(\sqrt{1 - 2 \cos(\beta_R) x + x^2} \right) - H_q \left(\sqrt{1 + 2 \cos(\beta_R) x + x^2} \right) \right) = 0 \quad (\text{A.2b})$$

$$b_R = \cos(\phi_R) H_q(1) \quad (\text{A.2c})$$

In such a case, $a_{ZR} = a_R + \beta_R$ is a zero response direction.

We now state the main lemma and apply it to prove Theorems A.1 and A.2. The remainder of the Appendix is the proof of the lemma.

Lemma A

Let ϕ_0, β_0 be given with $0 \leq \beta_0 \leq \frac{\pi}{2}$. Consider the following three conditions on a function $h(x)$:

(A1) $h(x)$ is even and continuous on the real line with $h(0) = 1$ and $h(\pm\infty) = 0$.

(A2) $\cos(\phi_0) \left(h \left(\sqrt{1 - 2x \cos(\beta_0) + x^2} \right) + h \left(\sqrt{1 + 2x \cos(\beta_0) + x^2} \right) - 2h(1)h(x) \right) = 0$ for all x .

(A3) $\sin(\phi_0) \left(h \left(\sqrt{1 - 2x \cos(\beta_0) + x^2} \right) - h \left(\sqrt{1 + 2x \cos(\beta_0) + x^2} \right) \right) = 0$ for all x .

Then:

(1a) Assume $\beta_0 = \frac{\pi}{2}$ and $\cos(\phi_0) = 0$. Then $h(x)$ satisfies the three conditions iff it satisfies (A1).

(1b) Assume $\beta_0 = \frac{\pi}{2}$ and $\cos(\phi_0) = 0$. Then:

(1b.1) $h(x)$ satisfies the three conditions with $h(1) > 0$ iff $h(x) = h(1)x^2 R(x^2)$ where $R(y)$ is $C[0, +\infty)$, $R(0) = 1$, $R(y+1) = R(y)$ for $y \geq 0$, and $0 < h(1) < 1$.

(1b.2) $h(x)$ satisfies the three conditions with $h(1) = 0$ iff $h(x) = 0$ for $|x| \geq 1$ with $h(x)$ even and continuous on the real line and $h(0) = 1$.

(1b.3) $h(x)$ satisfies the three conditions with $h(1) < 0$ iff $h(x) = (-h(1))x^2 R(x^2)$ where $R(y)$ is $C[0, +\infty)$, $R(0) = 1$, $R(y+1) = -R(y)$ for $y \geq 0$, and $-1 < h(1) < 0$.

(2a) Assume $0 < \beta_0 < \frac{\pi}{2}$ and $\cos(\phi_0) = \pm 1$. Then $h(x)$ satisfies the three conditions iff $h(x) = 0$ for $|x| \geq \sin(\beta_0)$ with $h(x)$ even and continuous on the real line and $h(0) = 1$.

(2b) Assume $0 < \beta_0 < \frac{\pi}{2}$ and $\cos(\phi_0) = \pm 1$. Then:

(2b.1) If $h(1) = 0$, then $h(x)$ satisfies the three conditions iff $h(x) = 0$ for $|x| \geq \sin(\beta_0)$ with $h(x)$ even and continuous on the real line and $h(0) = 1$.

(2b.2) If $0 < |h(1)| < 1$, then $h(x)$ satisfies the three conditions iff $h(x)$ satisfies (A1) and (A2).

(2b.3) If $|h(1)| = 1$, then no $h(x)$ satisfies the three conditions.

(3) Assume $\beta_0 = 0$. Then no $h(x)$ satisfies the three conditions.

Proof of Theorem A.1

Assume $R\left(\frac{-x}{x}\right)$ has the ZRD Property with a ZRD α_{ZR} . Apply Lemma 2' with $\beta_R = \alpha_{ZR} - \alpha_R$. Then the Hankel transform, extended to the real line as an even function $H_q(x)$, satisfies hypothesis (A1) of Lemma A and satisfies the following relations for all x , corresponding to hypotheses (A2) and (A3) of Lemma A:

$$\cos(\phi_R) \left(H_q \left(\sqrt{1 - 2 \cos(\beta_R)} x + x^2 \right) + H_q \left(\sqrt{1 + 2 \cos(\beta_R)} x + x^2 \right) - 2 H_q(1) H_q(x) \right) = 0 \quad (\text{A.3a})$$

$$\sin(\phi_R) \left(H_q \left(\sqrt{1 - 2 \cos(\beta_R)} x + x^2 \right) - H_q \left(\sqrt{1 + 2 \cos(\beta_R)} x + x^2 \right) \right) = 0 \quad (\text{A.3b})$$

Apply Lemma A. Since $\cos(\phi_R) = 0$, conclusion (1a) of Lemma A does not occur. By conclusion (3) of Lemma A, $\beta_R = 0$ does not occur, so $0 < \beta_R < \pi/2$.

Assume $0 < \beta_R < \pi/2$ occurs. Then either conclusion (2a) or (2b) of Lemma A occurs. Assume conclusion (2a) occurs. Then $H_q(x)$ satisfies $H_q(x) = 0$ for $|x| \geq \sin(\beta_R)$, where we already know that $\alpha_{ZR} = \alpha_R + \beta_R$ is a ZRD. Notice that, for each β_R' with $\beta_R < \beta_R' < \pi/2$, (A.3a) and (A.3b) are still satisfied for $H_q(x)$, hence $\alpha_{ZR}' = \alpha_R + \beta_R'$ is also a ZRD by Lemma 2', giving a sector of ZRDs. Result (3) of Theorem A.1 follows by extending the interval where $H_q(x)$ vanishes to its maximum extent. Now assume conclusion (2b) of Lemma A occurs. By (2b.3), $|H_q(1)| = 1$ does not occur. Then either $H_q(1) = 0$, in which case (2b.2) occurs and gives result (4) of Theorem A.1, or $H_q(1) \neq 0$ occurs, in which case (2b.1) occurs and gives result (3) of Theorem A.1 (after repeating the analysis for conclusion (2a)).

We can now assume (A.3a) and (A.3b) hold for $\beta_R = \pi/2$ and for no other value of β_R (otherwise, we return to the previous case). Since (A.3a) and (A.3b) hold only for this value, Lemma 2' implies the ZRD is the unique value $\alpha_{ZR} = \alpha_R + \pi/2$. Only conclusion (1b) of Lemma A applies, and its subcases (1b.1,2,3) give results (1), (2), (3) of Theorem A.1.

Conversely, assume a receptive field function $R\left(\frac{-}{x}\right)$ is given such that (3.1) and (3.2) hold and let the resulting Hankel transform $H_q(\rho)$ be extended to the real line as an even function. We wish to show that the field function has the ZRD Property.

Assume result (1) of Theorem A.1 holds, so that $\alpha_{ZR} = \alpha_R + \pi/2$ is a ZRD. Apply Lemma 2' with $\beta_R = \alpha_{ZR} - \alpha_R$ and b_R defined by (A.2c). Then direct substitution shows that (A.2a) and (A.2b) are satisfied, so the ZRD Property holds. Similarly, by direct substitution into the conditions of Lemma 2', results (2), (3), (4) of Theorem A.1 also imply the ZRD Property, noting that results (3) and (4), although stated for $\rho \neq 0$, are preserved for even extensions of $H_q(\rho)$ to the real line.

Theorem A.1 is proved.

Proof of Theorem A.2

A receptive field function $R\left(\frac{-}{x}\right)$ is given such that (3.1) and (3.2) hold. Let the Hankel transform be extended to the real line as an even function $H_q(x)$. We wish to determine the ZRDs (if any). By Lemma 2', each ZRD α_{ZR} determines a corresponding $\beta_R = \alpha_{ZR} - \alpha_R$ for which conditions (A.2a,b,c) are satisfied. That is, $H_q(x)$ necessarily satisfies hypothesis (A1) of Lemma A and conditions (A.2a,b) correspond to hypotheses (A2) and (A3) of Lemma A. We can now apply Lemma A to determine the values β_R . Since $\cos(\phi_R) = 0$, only conclusions (1a), (2a), and (3) of Lemma A can occur. By conclusion (3), $\beta_R = 0$ does not occur, so $0 < \beta_R < \pi/2$.

Assume $\beta_R = \pi/2$. Then conclusion (1a) of Lemma A occurs, and conditions (A.2a,b) are automatically satisfied because they are vacuous. That is, there is always a ZRD, namely, $\alpha_{ZR} = \alpha_R + \pi/2$, proving result (1) of Theorem A.2.

Assume $0 < \beta_R < \pi/2$. Then conclusion (2a) of Lemma A occurs, and $H_q(x)$ satisfies $H_q(x) = 0$ for $|x| > \sin(\beta_R)$. We can therefore form equations (A.2a,b) not only for β_R but for each β_R' with $\beta_R < \beta_R' < \pi/2$ as well. By Lemma 2', there is a sector of ZRDs, namely, $\alpha_{ZR}' = \alpha_R + \beta_R'$.

We have now shown that there is always at least one ZRD, corresponding to $\beta_R = \pi/2$, and that if there is more than one ZRD, the ZRDs must occur as a sector, corresponding to $\beta_R < \beta_R' < \pi/2$ (and thus including $\pi/2$). Furthermore, we have shown that if more than one ZRD occurs, then $H_q(x)$ must satisfy $H_q(x) = 0$ for $|x| > s_0$ for some $0 < s_0 < 1$ and that, when $H_q(x)$ satisfies such a condition, a sector of ZRDs is implied. Altogether, this proves result (2) of Theorem A.2.

Theorem A.2 is proved.

Proof of Lemma A

(1a) Assume $\beta_0 = \frac{\pi}{2}$ and $\cos(\phi_0) = 0$. Then (A2) and (A3) are vacuous. Hence $h(x)$ satisfies the three conditions iff it satisfies (A1).

(1b) Assume $\beta_0 = \frac{\pi}{2}$ and $\cos(\phi_0) = 0$. Notice (A3) is vacuous, and (A2) reduces to $h\left(\sqrt{1+x^2}\right) = h(1)h(x)$ for all x .

If $h(1) = 0$, it follows that $h(x) = 0$ for $|x| \geq 1$. That is, (A1,2,3) and $h(1) = 0$ give the conclusion of (1b.2), and the converse follows directly.

If $h(1) > 0$, define $F(y) := h(1)^{-y}h(\sqrt{y})$ for $y \geq 0$. Then $F(y)$ is $C[0, +\infty)$, $F(0) = 1$, and the reduced form of (A2) gives $F(y+1) = F(y)$. Then $h(x) = h(1)^{x^2}F(x^2)$. Using $h(\pm\infty) = 0$ from (A1) gives $0 < h(1) < 1$. This proves the representation (1b.1) for $h(x)$, and the converse follows directly.

If $h(1) < 0$, define $F(y) := |h(1)|^{-y}h(\sqrt{y})$ for $y \geq 0$. Proceeding as for $h(1) > 0$ proves the representation (1b.3), and the converse follows directly.

(2a) Assume $0 < \beta_0 < \frac{\pi}{2}$ and $\cos(\phi_0) = \pm 1$. Then, since $\sin(\phi_0) = 0$, (A3) implies

$$h\left(\sqrt{\sin(\beta_0)^2 + (x + \cos(\beta_0))^2}\right) = h\left(\sqrt{\sin(\beta_0)^2 + (x - \cos(\beta_0))^2}\right) \text{ for all } x$$

The function $h\left(\sqrt{\sin(\beta_0)^2 + x^2}\right)$ is therefore periodic and, by $h(\pm\infty) = 0$ from (A1), must be identically zero. Thus, $h(x) = 0$ for $|x| \geq \sin(\beta_0)$. In particular, $h(1) = 0$, so (A2) is necessarily satisfied whatever the value of $\cos(\phi_0)$. Representation (2a) follows. The converse is direct.

(2b) Assume $0 < \beta_0 < \frac{\pi}{2}$ and $\cos(\phi_0) = \pm 1$. Then $\sin(\phi_0) = 0$ and (A3) is vacuous.

If $h(1) = 0$, then (A2) reduces to

$$h\left(\sqrt{\sin(\beta_0)^2 + (x + \cos(\beta_0))^2}\right) = -h\left(\sqrt{\sin(\beta_0)^2 + (x - \cos(\beta_0))^2}\right) \text{ for all } x$$

The function $h\left(\sqrt{\sin(\beta_0)^2 + x^2}\right)$ is therefore periodic with period $4\cos(\beta_0)$ and, by $h(\pm\infty) = 0$ from (A1), must be identically zero. Thus, $h(x) = 0$ for $|x| \geq \sin(\beta_0)$. Representation (2b.1) follows, and the converse is direct.

If $0 < |h(1)| < 1$, then the stated result (2b.2) is immediate.

For $|h(1)| = 1$, first consider $h(1) = 1$. Notice (A2) becomes

$$\frac{1}{2} \left(h\left(\sqrt{\sin(\beta_0)^2 + (x - \cos(\beta_0))^2}\right) + h\left(\sqrt{\sin(\beta_0)^2 + (x + \cos(\beta_0))^2}\right) \right) = h(1)h(x) \text{ for all } x$$

By (A1), $h(x)$ must attain an absolute maximum $M = h(x_0) = h(1) = 1$. Applying (A2) with $x = x_0$ forces

$$h\left(\sqrt{\sin(\beta_0)^2+(x_0-\cos(\beta_0))^2}\right)=h\left(\sqrt{\sin(\beta_0)^2+(x_0+\cos(\beta_0))^2}\right)=M \text{ and } h(1)=1$$

Thus, $h(x)$ also attains the absolute maximum M at

$x_1=\sqrt{\sin(\beta_0)^2+(x_0+\cos(\beta_0))^2}>x_0+\cos(\beta_0)$. Repeating the argument at x_1 gives an absolute maximum at $x_2=\sqrt{\sin(\beta_0)^2+(x_1+\cos(\beta_0))^2}>x_0+2\cos(\beta_0)$, and so on, giving a sequence of points $x_n \rightarrow +\infty$ with $h(x_n)=M-1$, contradicting $h(\pm\infty)=0$ in (A1).

Now consider $h(1) = -1$. By (A1), $h(x)$ must attain an absolute maximum M $h(0)=1$, and $h(x)$ must have an absolute minimum m $h(1) = -1$. Assume $M = -m$. Let $h(x_0)=M$. Applying (A2) with $x = x_0$ gives

$$\frac{1}{2}\left(h\left(\sqrt{\sin(\beta_0)^2+(x_0-\cos(\beta_0))^2}\right)+h\left(\sqrt{\sin(\beta_0)^2+(x_0+\cos(\beta_0))^2}\right)\right)=h(1)h(x_0)\leq -M\leq m$$

This forces $m = -M$ and $h(1) = -1$ and

$$h\left(\sqrt{\sin(\beta_0)^2+(x_0-\cos(\beta_0))^2}\right)=h\left(\sqrt{\sin(\beta_0)^2+(x_0+\cos(\beta_0))^2}\right)=m$$

Thus, $h(x)$ attains the absolute minimum m at $x_{1/2}=\sqrt{\sin(\beta_0)^2+(x_0+\cos(\beta_0))^2}>x_0+\cos(\beta_0)$. Repeating the argument at $x = x_{1/2}$ gives

$$\frac{1}{2}\left(h\left(\sqrt{\sin(\beta_0)^2+(x_{1/2}-\cos(\beta_0))^2}\right)+h\left(\sqrt{\sin(\beta_0)^2+(x_{1/2}+\cos(\beta_0))^2}\right)\right)=h(1)h(x_{1/2})=-m=M$$

forcing $h(x_1)=M$ at $x_1=\sqrt{\sin(\beta_0)^2+(x_{1/2}+\cos(\beta_0))^2}>x_0+2\cos(\beta_0)$. Repeat to construct a sequence $x_n \rightarrow +\infty$ such that $h(x_n)=M-1$, contradicting $h(\pm\infty)=0$ in (A1). Now assume $M < -m$. Let $h(x_0)=m$. Applying (A2) with $x = x_0$ gives

$$\frac{1}{2}\left(h\left(\sqrt{\sin(\beta_0)^2+(x_0-\cos(\beta_0))^2}\right)+h\left(\sqrt{\sin(\beta_0)^2+(x_0+\cos(\beta_0))^2}\right)\right)=h(1)h(x_0)\geq -m>M$$

But $h\left(\sqrt{\sin(\beta_0)^2+(x\pm\cos(\beta_0))^2}\right)\leq M$, contradiction. This proves (2b.3).

(3) Assume $\beta_0 = 0$. Then $\sqrt{1\pm 2x\cos(\beta_0)+x^2}=|1\pm x|$. If $\sin(\phi_0) \neq 0$, then (A3) implies $h(x)$ is periodic, forcing a contradiction by the argument of case (2a). If $\sin(\phi_0) = 0$, then, since $h(x)$ is even, (A2) becomes an ordinary difference equation, namely, $h(x-1)+h(x+1)=2h(1)h(x)$. If $|h(1)| > 1$, there are two real characteristic roots with one root greater than one. If $|h(1)| = 1$, there is a double root, either $+1$ or -1 . If $|h(1)| < 1$, there are two complex conjugate roots with modulus one. Picking arbitrary starting points x_0, x_0+1 , it can be

shown by explicit solution that, for all these cases, $h(\pm\infty) = 0$ holds iff $h(x_0) = h(x_0 + 1) = 0$, forcing $h(x) = 0$ for all x and thus contradicting (A1).

Lemma A is proved.

Appendix B. Proof for Theorem B.1

This appendix proves Theorems B.1, the characterization theorem for elementary receptive field functions with the MRD Property.

The following lemma is used in the proof of Theorem B.1.

Lemma B

The following are equivalent:

1. $h(x)$ is $C[0, +\infty)$ and satisfies $h\left(\sqrt{1+x^2}\right) = h(1)h(x)$ for $x \geq 0$ with $h(1) \neq 0$.
2. $h(x) = e^{-cx^2} f(x^2)$ where $f(y)$ is $C[0, +\infty)$, $f(1) = \pm 1$, and $f(y+1) = f(1)f(y)$ for $y \geq 0$.

In case (1), $h(0) = 1$ necessarily holds. In case (2), $f(0) = 1$ necessarily holds.

Proof of Lemma B

Assume (1) holds. Define $f(y) := e^{cy} h(\sqrt{y})$ for $y \geq 0$, where $e^{-c} = |h(1)| > 0$. Then $f(y)$ is $C[0, +\infty)$, $f(1) = h(1) / |h(1)| = \pm 1$, and $f(y+1) = f(1)f(y)$ follows using the recursion relation in (1). Assume (2) holds. Then $h(x)$ is $C[0, +\infty)$, $h(1) = e^{-c} |f(1)| \neq 0$, and $h\left(\sqrt{1+x^2}\right) = h(1)h(x)$ follows by substitution. The lemma is proved.

Proof of Theorem B.1

It may be helpful to outline the structure of the proof. We first assume $R\left(\frac{\cdot}{x}\right)$ has the MRD Property. Then, as noted in Section 2, it has the ZRD Property and, since $\cos(\phi_R) \neq 0$, Theorem A.1 applies. We consider the four cases of Theorem A.1 in reverse order:

- *case (4), which will be shown inconsistent with the MRD Property;
- *case (3), which will imply result (2) of Theorem B.1;
- *cases (1) and (2) together, where case (2) will be shown inconsistent with the MRD Property and case (1) will imply result (1) of Theorem B.1.

Having derived cases (1) and (2) of Theorem B.1, we will then prove the converse, that each of these cases implies the MRD Property.

Assume $R\left(\frac{\cdot}{x}\right)$ has the MRD Property

Then $R\left(\frac{\cdot}{x}\right)$ has the ZRD Property, Theorem A.1 applies, and we will work through its four cases.

Assume case (4) of Theorem A.1 holds

Then $\cos(\phi_R) = \pm 1$ and the Hankel transform $H_q(\rho)$ satisfies

$$H_q\left(\sqrt{1-2\rho\cos(\zeta_0)+\rho^2}\right)+H_q\left(\sqrt{1+2\rho\cos(\zeta_0)+\rho^2}\right)=2H_q(1)H_q(\rho) \text{ for } \rho \geq 0$$

with $0 < |H_q(1)| < 1$, where ζ_0 is some fixed value with $0 < \zeta_0 \leq \frac{\pi}{2}$, and where ZRDs are given by $\alpha_{ZR} - \alpha_R = \pm \zeta_0$.

The MRD Property then implies that there must be a sector of ZRDs, $|\alpha_{ZR} - \alpha_R| = \zeta_0$. Consequently, applying Lemma 2 with $\cos(\phi_R) = \pm 1$ yields, for $\rho > 0$,

$$H_q\left(\sqrt{1-2\rho\cos(\alpha)+\rho^2}\right)+H_q\left(\sqrt{1+2\rho\cos(\alpha)+\rho^2}\right)=2H_q(1)H_q(\rho) \text{ for } \zeta_0 \leq \alpha \leq \frac{\pi}{2} \quad (\text{B.1})$$

Taking $\alpha = \frac{\pi}{2}$ and using the fact that $H_q(\rho)$ is $C[0, \infty)$ gives

$$H_q\left(\sqrt{1+\rho^2}\right)=H_q(1)H_q(\rho) \text{ for } \rho \geq 0 \quad (\text{B.2})$$

Since $H_q(1) \neq 0$, Lemma B applies to give $H_q(\rho) = e^{-c\rho^2} f(\rho^2)$ for $\rho \geq 0$ with corresponding conditions on $f(y)$. Since $f(y)$ is already periodic (with period 2) on the half-line, it can be extended to the whole line and the extended function will be useful below. So the representation can be summarized as

$$H_q(\rho) = e^{-c\rho^2} f(\rho^2) \text{ for } \rho \geq 0 \text{ where } f(y) \text{ is } C(-\infty, +\infty), f(1) = \pm 1, \text{ and } f(y+1) = f(1)f(y) \text{ for all } y. \quad (\text{B.3})$$

Notice $c > 0$ because $H_q(+\infty) = 0$. Combining this representation with (B.1) gives, for each $\rho > 0$,

$$e^{+2c\cos(\alpha)\rho} f\left(1-2\rho\cos(\alpha)+\rho^2\right) + e^{-2c\cos(\alpha)\rho} f\left(1+2\rho\cos(\alpha)+\rho^2\right) = 2f(1)f(\rho^2) \text{ for } \zeta_0 \leq \alpha \leq \frac{\pi}{2} \quad (\text{B.4a})$$

Setting $x = 2\cos(\alpha)\rho$ and using the symmetry of the expression gives, for each $\rho > 0$,

$$e^{+cx} f(1-x+\rho^2) + e^{-cx} f(1+x+\rho^2) = 2f(1)f(\rho^2) \text{ for } |x| \leq 2\cos(\zeta_0)\rho \quad (\text{B.4b})$$

Setting $\rho = 2N$, where N is a positive integer, and using the properties of f gives

$$e^{+cx} f(-x) + e^{-cx} f(x) = 2 \text{ for } |x| \leq 4\cos(\zeta_0)N \quad (\text{B.4c})$$

and letting $N \rightarrow \infty$ gives

$$e^{+cx} f(-x) + e^{-cx} f(x) = 2 \text{ for } x \in \mathbb{R} \quad (\text{B.4d})$$

However, dividing by e^{+cx} and taking a limit, implies $f(-\infty) = 0$, a contradiction since f is periodic and $f(0) = 1$. Consequently, case (4) of Theorem A.1 cannot occur.

Assume case (3) of Theorem A.1 holds

The Hankel transform $H_q(\rho)$ is then zero on an interval of the form $0 < \sin(\zeta_R) \rho < +\infty$, the largest such interval on which $H_q(\rho)$ vanishes. In particular, $H_q(\rho) = 0$ for $\rho \geq 1$, which implies, writing $a = a_p - a_R$,

$$b_R = \cos(\phi_R) H_q(1) = 0 \quad \text{and} \quad H_q\left(\sqrt{1+2\cos(\alpha)} \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) = 0 \quad \text{for} \quad 0 \leq \alpha \leq \frac{\pi}{2} \quad (\text{B.5})$$

Consequently, (3.9a) simplifies to

$$N(\alpha_p, \lambda_p) = \left| H_q\left(\sqrt{1+2\cos(\alpha)} \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}\right) \right| \quad \text{for} \quad 0 \leq \alpha \leq \frac{\pi}{2} \quad (\text{B.6})$$

which becomes, for $\lambda_p = \lambda_R$,

$$N(\alpha_p, \lambda_R) = \left| H_q\left(\sqrt{2-2\cos(\alpha)}\right) \right| \quad \text{for} \quad 0 \leq \alpha \leq \frac{\pi}{2} \quad (\text{B.7})$$

Then the MRD Property says that N must strictly decrease from its maximum $|H_q(0)|$ (at $a = 0$) until it reaches zero (at $2 - 2\cos(\alpha) = \sin(\zeta_R)^2$) and then remains zero (as a increases to $\pi/2$). Thus, $H_q(\rho)$ must strictly decrease from its maximum of one at $\rho = 0$ to zero at $\rho = \sin(\zeta_R)$, in particular, $H_q(\rho) = 0$. This proves that the condition of result (2) of Theorem B.1 is necessary for the MRD Property. Observation (a), that the MRD is given by $a_{MR} = a_R$, is a general result noted in the discussion of Lemma 3, and observation (b) on the sector of ZRDs follows from Theorem A.1.

Assume case (1) or case (2) of Theorem A.1 holds

Then in both cases the Hankel transform can be written as $H_q(\rho) = e^{-c\rho^2} F(\rho^2)$, where $e^{-c} = |H_q(1)|$ with $c > 0$ and $F(y)$ is periodic with period 2 on the half-line. The periodic extension of $F(y)$ to the whole line will be useful, and the representation can be summarized as

$$H_q(\rho) = e^{-c\rho^2} F(\rho^2) \quad \text{for all } \rho \geq 0 \quad \text{where } F(y) \text{ is } C(-\infty, +\infty), F(0) = 1, F(1) = \pm 1, \text{ and } F(y+1) = F(1)F(y) \quad (\text{B.8})$$

where $F(1) = +1$ corresponds to case (1) and $F(1) = -1$ to case (2) of Theorem A.1.

Combining this representation with (3.9b) and writing $a = a_p - a_R$ and $r = \lambda_R / \lambda_p$ gives

$$N(\alpha_p, \lambda_p)^2 = e^{-c(1+r^2)} \cdot N_1(\alpha, r)^2 \quad (\text{B.9a})$$

where $N_1(\alpha, r) > 0$ is given by

$$N_1(\alpha, r)^2 := \cos(\phi_R)^2 \left(e^{-2c\cos(\alpha)r} F(1+2\cos(\alpha)r+r^2) + e^{+2c\cos(\alpha)r} F(1-2\cos(\alpha)r+r^2) - 2F(1)F(r^2) \right)^2 + \sin(\phi_R)^2 \left(e^{-2c\cos(\alpha)r} F(1+2\cos(\alpha)r+r^2) - e^{+2c\cos(\alpha)r} F(1-2\cos(\alpha)r+r^2) \right)^2 \quad (\text{B.9b})$$

The MRD Property implies that, for each $r > 0$, $N_1(\alpha, r)$ is either identically zero or is initially positive and then strictly decreasing to zero (and remains zero) as a increases from 0 to $\pi/2$. Set $y = 2\cos(\alpha)r$ and notice the resulting expression for N_1 is even in y . Then, for each $r > 0$ and for $|y| \leq 2r$,

$$N_1^2 = \frac{\cos(\phi_R)^2 \left(e^{-cy} F(1+y+r^2) + e^{+cy} F(1-y+r^2) - 2F(1)F(r^2) \right)^2 + \sin(\phi_R)^2 \left(e^{-cy} F(1+y+r^2) - e^{+cy} F(1-y+r^2) \right)^2}{\sin(\phi_R)^2 \left(e^{-cy} F(1+y+r^2) - e^{+cy} F(1-y+r^2) \right)^2} \quad (\text{B.9c})$$

where N_1 is either zero on the whole interval $|y| \leq 2r$ or is initially zero and then strictly increasing as $|y|$ increases to $2r$. Setting $r^2 = a + 2N$, where a is an arbitrary real value and N is a sufficiently large positive integer, and using the properties of F from (B.8) gives, for $|y| \leq 2\sqrt{a+2N}$,

$$N_1^2 = \cos(\phi_R)^2 \left(e^{-cy} F(a+y) + e^{+cy} F(a-y) - 2F(a) \right)^2 + \sin(\phi_R)^2 \left(e^{-cy} F(a+y) - e^{+cy} F(a-y) \right)^2 \quad (\text{B.9d})$$

Let $N \rightarrow +\infty$. Then, for each real a , the expression T_F defined by

$$T_F(\phi_R, c, a; y) := \frac{\cos(\phi_R)^2 \left(e^{cy} F(a-y) + e^{-cy} F(a+y) - 2F(a) \right)^2 + \sin(\phi_R)^2 \left(e^{cy} F(a-y) - e^{-cy} F(a+y) \right)^2}{\sin(\phi_R)^2 \left(e^{cy} F(a-y) - e^{-cy} F(a+y) \right)^2} \quad (\text{B.9e})$$

is either zero for all $|y|$ or is initially zero and then strictly increasing as $|y|$ increases indefinitely. In particular, $T_F(y)$ is nondecreasing on $y \geq 0$.

CLAIM 1. $e^{2cy} F(-y)^2$ is nondecreasing on $-\infty < y < +\infty$.

PROOF. Let $a = 0$ and $y = x + 2N$ for positive integer N in (B.9e) with $0 \leq x \leq 2$. Then

$$e^{-4Nc} T_F(x+2N) = \frac{\cos(\phi_R)^2 \left(e^{cx} F(-x) + e^{-cx-4Nc} F(x) - 2e^{-2Nc} F(a) \right)^2 + \sin(\phi_R)^2 \left(e^{cx} F(-x) - e^{-cx-4Nc} F(x) \right)^2}{\sin(\phi_R)^2 \left(e^{cx} F(-x) - e^{-cx-4Nc} F(x) \right)^2} \quad (\text{B.10})$$

is a sequence of nondecreasing functions converging uniformly on $0 \leq x \leq 2$ to $(e^{cx} F(-x))^2$. The limit function is then nondecreasing on $0 \leq x \leq 2$ and, since shifting by 2 simply multiplies the function by a constant, that is, $(e^{c(x+2)} F(-(x+2)))^2 = e^{4c} (e^{cx} F(-x))^2$, the function must be nondecreasing on the entire line.

CLAIM 2. $e^{cy} F(-y)$ is positive and nondecreasing on $-\infty < y < +\infty$.

PROOF. If this function was zero at some value $y = y_0$, then $e^{2cy} F(-y)^2$ would, by Claim 1, be zero on $y \geq y_0$ and, by periodicity of F , be identically zero. But it has the value 1 at $y = 0$. Consequently, $e^{cy} F(-y)$ is nonzero and thus does not change sign and is necessarily positive. It is then nondecreasing since its square is nondecreasing by Claim 1.

CLAIM 3. Case (2) of Theorem A.1 does not hold.

PROOF. In case (2), $F(y)$ satisfies $F(y+1) = -F(y)$. Thus, $F(0) = 1$ and $F(1) = -1$, contradicting Claim 2.

As a result of Claim 3, only case (1) of Theorem A.1 holds, which implies that $F(y)$ is periodic with period 1. This proves the initial part of result (1) of Theorem B.1.

CLAIM 4. For each real a , $T_F(\phi_R, c, a; y)$ is nondecreasing on $0 \leq y \leq \delta_a$ and is strictly increasing on $\delta_a < y < +\infty$, where $0 \leq \delta_a < 1$.

PROOF. As noted in connection with (B.9e), $T_F(y)$ is either (a) zero for all $y \geq 0$ or (b) initially zero and then strictly increasing. Since $F(y)$ has period 1 and is nonzero, we have

$$T_F(1) = ((\cos(\phi_R))^2(e^c + e^{-c} - 2)^2 + (\sin(\phi_R))^2(e^c - e^{-c})^2) F(a)^2 > 0$$

Alternative (b) must hold, and $T_H(y)$ is in the strictly increasing regime at $y = 1$, so $\delta_a < 1$.

CLAIM 5. $e^{cy}F(-y)$ is positive and strictly increasing on $-\infty < y < +\infty$.

PROOF. By Claim 2, the function is positive and nondecreasing. Assume it is not strictly increasing. Then it must be constant on some interval:

$$e^{+cx}F(-x) = e^{+cb}F(-b) \quad \text{on} \quad |x - b| < \delta \quad (\text{B.11a})$$

which implies

$$e^{-cx}F(+x) = e^{+cb}F(-b) \quad \text{on} \quad |x + b| < \delta \quad (\text{B.11b})$$

Consider Claim 4 with $a = -b$ and $y = z + N$ for integers $N \gg 1$ and $|z| < \delta$:

$$\begin{aligned} T_F(\phi_R, c, -b; y) &= (\cos(\phi_R))^2 (e^{c(z+N)}F(-b-z) + e^{-c(z+N)}F(-b+z) - 2F(-b))^2 + \\ &\quad (\sin(\phi_R))^2 (e^{c(z+N)}F(-b-z) - e^{-c(z+N)}F(-b+z))^2 = \\ &\quad ((\cos(\phi_R))^2 (e^{cN} + e^{-cN} - 2)^2 + (\sin(\phi_R))^2 (e^{cN} - e^{-cN})^2) F(-b)^2 \end{aligned}$$

That is, $T_H(y)$ is constant on subintervals $|y - M| < \delta$ for sufficiently large N , contradicting Claim 4.

Note that Claim 5 implies the functions $e^{cy}F(a-y) = e^{ca}e^{c(y-a)}F(a-y)$ are strictly increasing and the functions $e^{-cy}F(a+y)$ are strictly decreasing. In particular, Claim 5 gives (c) under result (1) of Theorem B.1.

CLAIM 6. Assume $\cos(\phi_R) \neq \pm 1$. Then, for each real a , $T_H(\phi_R, c, a, y)$ is strictly increasing on $0 < y < +\infty$.

PROOF. By Claim 4, for each a , $T_H(\phi_R, c, a, y)$ must be initially zero, then strictly increasing. We show that the initially zero interval must always reduce to a single point. Assume it does not. Then, for some a_0 , $T_H(a_0; y) = 0$ on some interval $|y| \leq \delta_0$ with $\delta_0 > 0$. Since $\sin(\phi_R) \neq 0$, $T_H(a_0; y) = 0$ gives two equations

$$e^{-cy}F(a_0+y) + e^{+cy}F(a_0-y) = 2F(a_0) \quad \text{and} \quad e^{-cy}F(a_0+y) - e^{+cy}F(a_0-y) = 0 \quad \text{on} \quad |y| \leq \delta_0$$

which can be solved to give $e^{-cy}F(a_0+y) = F(a_0)$ and $e^{+cy}F(a_0-y) = F(a_0)$ on $|y| \leq \delta_0$. Set $x = y + N$ for positive integers N with $|y| \leq \delta_0$ and combine these expressions with the periodicity of F to obtain

$$T_F(a_0; x) = F(a_0)^2 \left((\cos(\phi_R))^2 (e^{+cN} + e^{-cN} - 2)^2 + (\sin(\phi_R))^2 (e^{+cN} - e^{-cN})^2 \right)$$

This is a contradiction because it shows $T_H(a_0; x)$ to be constant on the intervals $|x - M| = |y| \delta_0$, but $T_H(a_0; x)$ must be strictly increasing on such intervals for sufficiently large N by Claim 4.

Claim 6 proves part (**) of result (1) for Theorem B.1.

CLAIM 7. Assume $\cos(\phi_R) = \pm 1$. Then:

(7a) For each a , the function $f(a; y) = e^{cy} F(a - y) + e^{-cy} F(a + y) - 2F(a)$ is initially zero on $0 \leq y \leq \delta_a$ and strictly increasing on $\delta_a \leq y < +\infty$. (The value δ_a satisfies the bound $0 \leq \delta_a < 1$.)

(7b) In (7a), $\delta_a > 0$ if and only if the value a satisfies $e^{-cy} F(y) = e^{-ca} F(a) - K(a)(y - a)$ on $|y - a| \leq \delta_a$ for some constant $K(a) > 0$.

PROOF. Since $\sin(\phi_R) = 0$, the even function $T_H(x, y)$ in (B.9e) reduces to $T_H(x, y) = f(x, y)^2$, which is nondecreasing for $y \geq 0$ by Claim 4. Since $f(x, 0) = 0$ and $f(x, y)$ is eventually positive due to the growth of the term $e^{+cy} F(a - y)$, we have $f(x, y) \geq 0$ and nondecreasing. Combining this behavior with the general behavior of $T_H(x, y)$ given by Claim 4, which also gives $\delta_a < 1$, proves Claim (7a).

For Claim (7b), it will be convenient to use the fact that $f(x, y)$ is an even function of y . Let a_0 be a value such that $f(a_0, y) = 0$ on $|y| \leq \delta_0$ for some positive δ_0 , that is,

$$e^{+cy} F(a_0 - y) + e^{-cy} F(a_0 + y) - 2F(a_0) = 0 \text{ on } |y| \leq \delta_0$$

This equation can be solved for e^{+cy} :

$$e^{+cy} = \frac{1}{F(a_0 - y)} \left(F(a_0) \pm \sqrt{F(a_0)^2 - F(a_0 + y) F(a_0 - y)} \right) \text{ for } |y| \leq \delta_0$$

that is,

$$e^{+cy} F(a_0 - y) = F(a_0) \left(1 \pm \sqrt{1 - K_0(y)} \right) \text{ for } |y| \leq \delta_0 \quad (\text{B.12})$$

where $K_0(y) = F(a_0 + y) F(a_0 - y) / F(a_0)^2$ is even, $K_0(0) = 1$, and $K_0(y) \leq 1$ for $|y| \leq \delta_0$. By Claim 5, $e^{+cy} F(a_0 - y)$ is strictly increasing, which forces

$$\begin{aligned} e^{+cy} F(a_0 - y) &= F(a_0) \left(1 + \sqrt{1 - K_0(y)} \right) \text{ for } 0 \leq y \leq \delta_0 \\ &= F(a_0) \left(1 - \sqrt{1 - K_0(y)} \right) \text{ for } -\delta_0 \leq y \leq 0 \end{aligned} \quad (\text{B.13})$$

with the even function $K_0(y)$ strictly decreasing as $|y|$ increases. Now consider $f(x, y)$ with $a = a_0 + \epsilon$, where $|\epsilon| \leq \delta_0 / 2$, which is an even function of y . Thus, for $|\epsilon| \leq \delta_0 / 2$ and $|y| \leq \delta_0 / 2$,

$$\begin{aligned} f(a_0 + \epsilon; y) &= e^{+cy} F(a_0 + \epsilon - y) + e^{-cy} F(a_0 + \epsilon + y) - 2F(a_0 + \epsilon) \\ &= e^{+c\epsilon} e^{+c(y-\epsilon)} F(a_0 - (y - \epsilon)) + e^{+c\epsilon} F(a_0 - (-y - \epsilon)) - 2e^{+c\epsilon} e^{c(-\epsilon)} e^{c(-y-\epsilon)} F(a_0 - (-\epsilon)) \\ &= e^{+c\epsilon} F(a_0) \left[\operatorname{sgn}(y - \epsilon) \sqrt{1 - K_0(y - \epsilon)} + \operatorname{sgn}(-y - \epsilon) \sqrt{1 - K_0(-y - \epsilon)} - 2 \operatorname{sgn}(-\epsilon) \sqrt{1 - K_0(-\epsilon)} \right] \\ &= e^{+c\epsilon} F(a_0) \left[\operatorname{sgn}(y - \epsilon) \sqrt{1 - K_0(y - \epsilon)} - \operatorname{sgn}(y + \epsilon) \sqrt{1 - K_0(y + \epsilon)} + 2 \operatorname{sgn}(\epsilon) \sqrt{1 - K_0(\epsilon)} \right] \end{aligned} \quad (\text{B.14})$$

Consider the two cases $\epsilon = \pm \epsilon_0$ with $\epsilon_0 > 0$. In each case set $y = \epsilon_0$ and use $f(x, y) \geq 0$ to obtain

$$f(a_0 + \epsilon_0; \epsilon_0) = e^{+c\epsilon_0} F(a_0) \left[0 - \sqrt{1 - K_0(2\epsilon_0)} + 2\sqrt{1 - K_0(\epsilon_0)} \right] \geq 0 \quad (\text{B.15a})$$

$$f(a_0 - \epsilon_0; \epsilon_0) = e^{-c\epsilon_0} F(a_0) \left[+\sqrt{1 - K_0(2\epsilon_0)} - 0 - 2\sqrt{1 - K_0(\epsilon_0)} \right] \geq 0 \quad (\text{B.15b})$$

Consequently,

$$\sqrt{1 - K_0(2\epsilon)} = 2\sqrt{1 - K_0(\epsilon)} \text{ for } |\epsilon| \leq \delta_0/2 \quad (\text{B.16})$$

Setting $z(\epsilon) = \sqrt{1 - K_0(\epsilon)}$, note that continuous solutions to the functional equation $z(2\epsilon) = 2z(\epsilon)$ on intervals $0 < \epsilon < \epsilon_0$ must be given by $z(\epsilon) = k_0\epsilon$ for constants k_0 . (Use the functional equation to derive this formula for a dense set and note a continuous function is determined by its values on a dense set.) Consequently, since $\sqrt{1 - K_0(\epsilon)}$ is a strictly increasing function of $|\epsilon|$,

$$\sqrt{1 - K_0(x)} = k_0|x| \text{ for some } k_0 > 0 \text{ and for } |x| \leq \delta_0 \quad (\text{B.17})$$

Eqn. (B.13) now becomes

$$e^{+cy} F(a_0 - y) = F(a_0) (1 + k_0 y) \text{ for some } k_0 > 0 \text{ and } |y| \leq \delta_0 \quad (\text{B.18a})$$

that is,

$$e^{-cy} F(y) = e^{-ca_0} F(a_0) - K_1(y - a_0) \text{ for some constant } K_1 > 0 \text{ and } |y - a_0| \leq \delta_0 \quad (\text{B.18b})$$

which is precisely the “only if” part of Claim (7 b).

To obtain the “if” part of Claim (7b), start with (B.18b), obtain expressions for $e^{-cy} F(a_0 + y)$ and $e^{+cy} F(a_0 - y)$, and observe that $f(a_0; y) = 0$ for $|y| \leq \delta_0$.

Claim 7 proves part (*) of result (1) of Theorem B.1, including the stated condition on the existence of nontrivial initially zero intervals. Altogether, Claims 1-7 complete the proof that the conditions in result (1) of Theorem B.1 are necessary for the MRD Property. Observation (a), that the MRD is given by $\alpha_{MR} = \alpha_R$, is a general result noted in the discussion of Lemma 3; observation (b) on the unique ZRD follows from Theorem A.1; and observation (c) was established by Claim 5.

To prove the converse

We wish to show that results (1) and (2) of Theorem B.1 are sufficient conditions, that is, that they imply the MRD Property. For each result, we calculate $N(\alpha_P, \lambda_P)$, given by (3.9b), and show that it satisfies the criteria of Lemma 3. Notice that the ZRD Property holds for each case, so the Balanced Field Property also holds and $b_R = \cos(\phi_R) H_q(1)$ in (3.9b) for both cases. Similarly, $\alpha_{MR} = \alpha_R$ in both cases.

Assume result (1) of Theorem B.1 holds

Then $H_q(\rho) = e^{-c\rho^2} F(\rho^2)$ with $c > 0$ and the stated conditions on $F(y)$ hold. Substituting into (3.9b) gives

$$N(\alpha_p, \lambda_p)^2 = \exp\left(-c\left(1 + \frac{\lambda_R^2}{\lambda_p^2}\right)\right) \cdot T_F\left(\phi_R, c, \frac{\lambda_R^2}{\lambda_p^2}; 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p}\right) \quad (\text{B.19})$$

where $T_F(\phi_R, c, a, y)$ is given by (B.9e) and we note the particular case $T_F(\pm \pi, c, a, y) = f(a, y)^2$. Since $T_F(\phi_R, c, a, y)$ is strictly increasing on $y \geq 0$ for arbitrary real a or (in the particular case $\phi_R = \pm \pi$) may be initially zero and then strictly increasing, it immediately follows that, for each fixed λ_p , N is strictly decreasing to zero as $\cos(\alpha_p - \alpha_R)$ decreases, that is, as $|\alpha_p - \alpha_R|$ increases, and either reaches zero at the ZRD $|\alpha_p - \alpha_R| = \pi/2$ or (in the particular case $\phi_R = \pm \pi$) may reach zero before then and remains zero until $|\alpha_p - \alpha_R| = \pi/2$. The conditions of Lemma 3 are satisfied.

Assume result (2) of Theorem B.1 holds

Then $H_q(\rho) = 0$ on $\sin(\zeta_R) \mid \rho < +\infty$ and is positive and strictly decreasing on $0 \leq \rho < \sin(\zeta_R)$, so $b_R = 0$ and (3.9b) reduce to the single term

$$N(\alpha_p, \lambda_p) = H_q\left(\sqrt{1 - 2 \cos(\alpha_p - \alpha_R) \frac{\lambda_R}{\lambda_p} + \frac{\lambda_R^2}{\lambda_p^2}}\right) \quad \text{for} \quad -\frac{\pi}{2} \leq \alpha_p - \alpha_R \leq +\frac{\pi}{2} \quad (\text{B.20})$$

It immediately follows that, for each fixed λ_p , N is strictly decreasing to zero, and remains zero, as $\cos(\alpha_p - \alpha_R)$ decreases, that is, as $|\alpha_p - \alpha_R|$ increases. The conditions of Lemma 3 are satisfied.

The proof of Theorem B.1 is now complete.

Appendix C. Miscellaneous Derivations

Section 3. Derivation of eqn. (3.6): Fourier transform of elementary receptive field function in terms of Hankel transforms

We wish to derive eqn. (3.6), the representation of the Fourier transform of an elementary receptive field function (3.1) in terms of the Hankel transform $H_q(\rho)$ of the weight function $q(r)$. By elementary properties of the Fourier transform (as defined by eqn. (2.4)), the Fourier transform of the elementary receptive field function $R(x_1, x_2)$ can be reduced to the Fourier transform of the weight function $p(x_1, x_2)$:

$$R(\vec{x}) = \frac{2\pi}{\lambda_R^2} q\left(2\pi \frac{|\vec{x}|}{\lambda_R}\right) \left(\cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) - b_R \right) = \frac{2\pi}{\lambda_R^2} p\left(\frac{2\pi}{\lambda_R} x_1, \frac{2\pi}{\lambda_R} x_2\right) \left(\cos\left(2\pi \frac{\vec{d}(\alpha_R) \cdot \vec{x}}{\lambda_R} - \phi_R\right) - b_R \right) \quad (\text{C.1a})$$

has the Fourier transform

$$F_R(s_1, s_2) = \frac{1}{2\pi} \left[\frac{1}{2} \left(e^{+i\phi_R} F_p\left(\frac{\lambda_R}{2\pi} \left(s_1 + \frac{1}{\lambda_R} \cos(\alpha_R)\right), \frac{\lambda_R}{2\pi} \left(s_2 + \frac{1}{\lambda_R} \sin(\alpha_R)\right)\right) + e^{+i\phi_R} F_p\left(\frac{\lambda_R}{2\pi} \left(s_1 + \frac{1}{\lambda_R} \cos(\alpha_R)\right), \frac{\lambda_R}{2\pi} \left(s_2 + \frac{1}{\lambda_R} \sin(\alpha_R)\right)\right) - b_R F_p\left(\frac{\lambda_R}{2\pi} s_1, \frac{\lambda_R}{2\pi} s_2\right) \right] \quad (\text{C.1b})$$

The 2D-Fourier transform of a circularly symmetric function can be expressed as the Hankel transform (defined by eqn. (3.4)) of the radial form:

$$p(x_1, x_2) = q(r) \quad (\text{C.1c})$$

has the Hankel transform

$$F_p(s_1, s_2) = 2\pi H_q(2\pi\rho) \quad \text{where} \quad \rho = \sqrt{s_1^2 + s_2^2} \quad (\text{C.1d})$$

Combining (C.1b) and (C.1d) gives (3.6).

Section 5a. Derivation of part (2) of the Closure Lemma for balanced Gabor weights

We wish to prove part (2) of the Closure Lemma for balanced Gabor weights. It is sufficient to prove the following:

CLAIM. Given $0 < \gamma_1 < \gamma_2$ and $G(y)$ continuous on the real line with $e^{-\gamma_1 y} G(y)$ strictly decreasing and positive and an arbitrary real a . If

$$e^{\gamma_1 y} G(a-y) + e^{-\gamma_1 y} G(a+y) \text{ is increasing for } y \geq 0, \quad (\text{C.2a})$$

then

$$e^{\gamma_2 y} G(a-y) + e^{-\gamma_2 y} G(a+y) \text{ is strictly increasing for } y > 0. \quad (\text{C.2b})$$

PROOF. Note $e^{-\gamma_1 y} G(y)$ is strictly decreasing and positive implies

$$e^{-\gamma_1 y} G(a+y) \text{ and } e^{-\gamma_2 y} G(a+y) \text{ are strictly decreasing and positive} \quad (\text{C.3a})$$

$$e^{+\gamma_1 y} G(a-y) \text{ and } e^{+\gamma_2 y} G(a-y) \text{ are strictly increasing and positive} \quad (\text{C.3b})$$

By hypothesis, for $0 < y_1 < y_2$,

$$e^{\gamma_1 y_2} G(a-y_2) + e^{-\gamma_1 y_2} G(a+y_2) \geq e^{\gamma_1 y_1} G(a-y_1) + e^{-\gamma_1 y_1} G(a+y_1) \quad (\text{C.4})$$

We will show

$$e^{\gamma_2 y_2} G(a-y_2) + e^{-\gamma_2 y_2} G(a+y_2) > e^{\gamma_2 y_1} G(a-y_1) + e^{-\gamma_2 y_1} G(a+y_1) \quad (*)$$

Multiply (C.4) by $e^{(\gamma_2 - \gamma_1)y_2}$ to obtain

$$e^{\gamma_2 y_2} G(a-y_2) + e^{(\gamma_2 - 2\gamma_1)y_2} G(a+y_2) \geq e^{\gamma_2 y_2} e^{\gamma_1(-y_2+y_1)} G(a-y_1) + e^{\gamma_2 y_2} e^{-\gamma_1(y_2+y_1)} G(a+y_1) \quad (\text{C.5a})$$

$$\begin{aligned} & e^{\gamma_2 y_2} G(a-y_2) + e^{-\gamma_2 y_2} G(a+y_2) - e^{\gamma_2 y_1} G(a-y_1) - e^{-\gamma_2 y_1} G(a+y_1) \geq \\ & + e^{-\gamma_2 y_2} G(a+y_2) - e^{(\gamma_2 - 2\gamma_1)y_2} G(a+y_2) + e^{\gamma_2 y_2} e^{\gamma_1(-y_2+y_1)} G(a-y_1) + e^{\gamma_2 y_2} e^{-\gamma_1(y_2+y_1)} G(a+y_1) \\ & - e^{-\gamma_2 y_1} G(a-y_1) - e^{-\gamma_2 y_1} G(a+y_1) \end{aligned} \quad (\text{C.5b})$$

The desired result (*) follows if we can show the right side $L > 0$. We have

$$\begin{aligned} L = & - \left(e^{2(\gamma_2 - \gamma_1)y_2} - 1 \right) e^{-\gamma_2 y_2} G(a+y_2) + \\ & \left(e^{\gamma_2 y_2} e^{\gamma_1(-y_2+y_1)} - e^{\gamma_2 y_1} \right) G(a-y_1) + \left(e^{\gamma_2 y_2} e^{-\gamma_1(y_2+y_1)} - e^{-\gamma_2 y_1} \right) G(a+y_1) \end{aligned} \quad (\text{C.6a})$$

As noted, $0 < e^{-\gamma_2 y_2} G(a+y_2) < e^{-\gamma_2 y_1} G(a+y_1)$, so

$$\begin{aligned}
L &> -\left(e^{2(\gamma_2-\gamma_1)y_2} - 1\right) e^{-\gamma_2 y_1} G(a+y_1) + \\
&\quad \left(e^{\gamma_2 y_2} e^{\gamma_1(-y_2+y_1)} - e^{\gamma_2 y_1}\right) G(a-y_1) + \left(e^{\gamma_2 y_2} e^{-\gamma_1(y_2+y_1)} - e^{-\gamma_2 y_1}\right) G(a+y_1) \\
&= \left(e^{(\gamma_2-\gamma_1)(y_2+y_1)} - e^{2(\gamma_2-\gamma_1)y_2}\right) e^{-\gamma_2 y_1} G(a+y_1) + \left(e^{\gamma_2 y_2} e^{\gamma_1(-y_2+y_1)} - e^{\gamma_2 y_1}\right) G(a-y_1) \\
&= \left(e^{(\gamma_2-\gamma_1)(y_2-y_1)} - 1\right) e^{\gamma_2 y_1} G(a-y_1) + e^{(\gamma_2-\gamma_1)y_2} \left(e^{(\gamma_2-\gamma_1)y_1} - e^{(\gamma_2-\gamma_1)y_2}\right) e^{-\gamma_2 y_1} G(a+y_1) > 0
\end{aligned}
\tag{C.6b}$$

The claim is proved.

Section 5a. Derivation of eqn. (5a.13c): Condition determining balanced Gabor weights of order 1

Balanced Gabor weights of order 1 have Hankel transforms $H_g(\rho) = e^{-\gamma_R \rho^2} G(\rho^2)$ where

$G(y) := \frac{1+c_R \cos(2\pi y - \psi_R)}{1+c_R \cos(\psi_R)}$, where $G(y)$ must satisfy the conditions of the balanced Gabor definition. We will derive eqn. (5a.13c), a characterization of these conditions for the order 1 case.

Requirement (a): $H_g(\rho) = e^{-c\rho^2} F(\rho^2)$ where $F(y)$ is continuous and positive on the real line, $F(0) = 1$, $F(y+1) = F(y)$. Solution: Obviously $G(y)$ satisfies this requirement iff $|c_R| < 1$.

Requirement (b): $e^{-cy} F(y)$ is strictly decreasing on the real line, that is, $e^{-cy} F(y)$ (with equality at most at isolated points). Solution: $G(y)$ satisfies this requirement iff

$$c_R^2 \leq \frac{\gamma_R^2}{4\pi^2 + \gamma_R^2}.$$

Derivation (b). The condition can be restated as $\frac{G'(y)}{G(y)} \leq c$. Then

$$\begin{aligned}
\frac{G'(y)}{G(y)} &= \frac{-2\pi c_R \sin(2\pi y - \psi_R)}{1+c_R \cos(2\pi y - \psi_R)} \text{ has min-max values } \frac{\pm 2\pi c_R}{\sqrt{1-c_R^2}} \text{ (occurring at } \cos(2\pi y - \psi_R) = \\
&\quad -c_R), \text{ that is, max-value } \frac{2\pi|c_R|}{\sqrt{1-c_R^2}}. \text{ Take } \frac{2\pi|c_R|}{\sqrt{1-c_R^2}} = \gamma_R, \text{ or } c_R^2 \leq \frac{\gamma_R^2}{4\pi^2 + \gamma_R^2}.
\end{aligned}$$

Requirement (c): For each real a , $T_F(c, a; y) := e^{cy} F(a-y) + e^{-cy} F(a+y)$ is strictly increasing for $y \geq 0$. That is, taking a derivative and expanding: for each real a , $(F'(a-y) - cF(a-y)) - e^{-2cy}(F'(a+y) - cF(a+y)) \leq 0$ for $y \geq 0$ (with equality at most at

isolated points). Solution: $G(y)$ satisfies this requirement iff $|c_R| \leq \frac{\gamma_R^2}{\gamma_R^2 + 4\pi^2}$.

Derivation (c). Set $G_1(y) := 1 + c_R \cos(2\pi y - \psi_R)$. The equivalent question is: for what range of c is

$$(G_1'(a-y) - cG_1(a-y)) - e^{-2cy}(G_1'(a+y) - cG_1(a+y)) \leq 0 \text{ on } y \geq 0 \text{ for all real } a?$$

Notice

$$G_1'(y) - cG_1(y) = -2\pi c_R \sin(2\pi y - \psi_R) - c(1+c_R \cos(2\pi y - \psi_R))$$

The question is the range of c such that, for $y = 0$,

$$[-2\pi c_R \sin(2\pi(a-y) - \psi_R) - c(1+c_R \cos(2\pi(a-y) - \psi_R))] - e^{-2cy} [-2\pi c_R \sin(2\pi(a+y) - \psi_R) - c(1+c_R \cos(2\pi(a+y) - \psi_R))] \leq 0$$

Writing $C_1 = \cos(2\pi a - \psi_R)$, $S_1 = \sin(2\pi a - \psi_R)$, this statement is equivalent to

$$c_R C_1 \left[(1+e^{-2cy}) 2\pi \sin(2\pi y) - (1-e^{-2cy}) c \cos(2\pi y) \right] + \quad \text{for } y \geq 0 \\ c_R S_1 \left[-(1-e^{-2cy}) 2\pi \cos(2\pi y) - (1+e^{-2cy}) c \sin(2\pi y) \right] \leq c(1-e^{-2cy})$$

that is,

$$c_R (C_1 [\coth(cy) 2\pi \sin(2\pi y) - c \cos(2\pi y)] + S_1 [-2\pi \cos(2\pi y) - \coth(cy) c \sin(2\pi y)]) \leq c \text{ for } y > 0$$

The expression on the left (given $C_1, S_1 = \cos(2\pi a - \psi_R), \sin(2\pi a - \psi_R)$) has the precise upper bound over all a :

$$|c_R| \sqrt{(\coth(cy) 2\pi \sin(2\pi y) - c \cos(2\pi y))^2 + (2\pi \cos(2\pi y) + \coth(cy) c \sin(2\pi y))^2} \leq c \text{ for } y > 0$$

That is, for each $y > 0$, the left side is simply the amplitude at y as a varies of the preceding expression and will be attained for infinitely many values of a . Hence the inequality is necessary and sufficient to insure that the preceding inequalities are maintained at each $y > 0$. It can be rewritten as

$$(\cos(2\pi y)^2 + \coth(cy)^2 \sin(2\pi y)^2)(4\pi^2 + c^2) c_R^2 \leq c^2 \text{ for } y > 0$$

By the following claim, this inequality will hold if we have $\left(1 + \frac{4\pi^2}{c^2}\right)(4\pi^2 + c^2) c_R^2 \leq c^2$, that

is, $c_R^2 \leq \left(\frac{c^2}{4\pi^2 + c^2}\right)^2$ the stated result.

CLAIM. $f(y) := \cos(2\pi y)^2 + \coth(cy)^2 \sin(2\pi y)^2 \leq 1 + \frac{4\pi^2}{c^2}$ for $y = 0$.

PROOF. Note $f(0) = 1 + \frac{4\pi^2}{c^2}$. Note $f(y) = \cos(2\pi y)^2 + (1 + \operatorname{csch}(cy)^2) \sin(2\pi y)^2 = 1 +$

$(\operatorname{csch}(cy) \sin(2\pi y))^2$. Consider max/min for $y > 0$. In particular, consider $g(y) = \frac{\sin(2\pi y)}{\sinh(cy)}$ for $y > 0$ and

$$g'(y) = \frac{\sinh(cy)(2\pi \cos(2\pi y)) - \sin(2\pi y)c \cosh(cy)}{\sin(cy)^2} = (\text{positive term})(\tanh(cy)(2\pi \cos(2\pi y)) - \sin(2\pi y)c$$

Hence critical points for $y > 0$ satisfy $\tanh(cy) = \tan(2\pi y) \frac{c}{2\pi}$. Note

$$\operatorname{csch}(cy)^2 = \frac{1 - \tanh(cy)^2}{\tanh(cy)^2} = \frac{1 - \tan(2\pi y)^2 \left(\frac{c}{2\pi}\right)^2}{\tan(2\pi y)^2 \left(\frac{c}{2\pi}\right)^2} \quad (\text{at critical points}), \text{ so}$$

$$f(y) = 1 + \operatorname{csch}(cy)^2 \sin(2\pi y)^2 = 1 + \frac{\cos(2\pi y)^2 - \sin(2\pi y)^2 \left(\frac{c}{2\pi}\right)^2}{\left(\frac{c}{2\pi}\right)^2} = \cos(2\pi y)^2 + \frac{\cos(2\pi y)^2}{\left(\frac{c}{2\pi}\right)^2} = \cos(2\pi y)^2 \left(1 + \frac{4\pi^2}{c^2}\right)$$

(at critical points $y > 0$). Hence $f(y) \leq 1 + \frac{4\pi^2}{c^2}$ for $y \geq 0$. The claim is proved.

References

1. Abramowitz, M.; Stegun, IA. Handbook of Mathematical Functions. National Bureau of Standards; 1972. Applied Mathematics Series 55
2. Albright TD, Stoner GR. Contextual influences on visual processing. Annual Review of Neuroscience. 2002; 25:339–379.
3. Anderson SJ, Burr DC. Spatial summation properties of directionally selective mechanisms in human vision. Journal of the Optical Society of America, A. 1991; 8:1330–1339.
4. Daugman JG. Two-dimensional spectral analysis of cortical receptive field profiles. Vision Research. 1980; 20:847–856. [PubMed: 7467139]
5. Daugman JG. Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America, A. 1985; 2:1160–1169.
6. DeValois, RL.; Albrecht, DG.; Thorell, LG. Cortical cells: Bar and edge detectors, or spatial frequency filters?. In: Cool, SJ.; Smith, EL., editors. Frontiers in Visual Science. Springer-Verlag; New York: 1978. p. 544–556.
7. DeValois, RL.; DeValois, KK. Spatial Vision. Oxford University Press; New York: 1988.
8. Field DJ. Relations between the statistics of natural images and the response properties of cortical cells. Journal of the Optical Society of America. 1987; 4:2379–2394. [PubMed: 3430225]
9. Foley JM, Varadharajan S, Koh CC, Farias MCQ. Detection of gabor patterns of different sizes, shapes, phases and eccentricities. Vision Research. 2007; 47:85–107. [PubMed: 17078992]
10. Gabor D. Theory of communication. J. Inst. Electrical Engineers. 1946; 93:429–457.
11. Heggelund P. Quantitative studies of the discharge fields of single cells in cat striate cortex. J. Physiol. Lond. 1986a; 373:277–292. [PubMed: 3746674]
12. Heggelund P. Quantitative studies of enhancement and suppression zones in the receptive field of simple cells in cat striate cortex. J. Physiol. Lond. 1986b; 373:293–310. [PubMed: 3746675]
13. Heitger F, Rosenthaler L, von der Heydt R, Peterhans E, Kubler O. Simulation of neural contour mechanisms: From simple to end-stopped cells. Vision Research. 1992; 32:963–981. [PubMed: 1604865]
14. Hubel DH, Weisel TN. Receptive fields and functional architecture of monkey striate cortex. J. Physiol. Lond. 1968; 195:215–243. [PubMed: 4966457]
15. Hubel DH, Weisel TN. Functional architecture of the macaque monkey visual cortex. Proc. Roy. Soc. Lond., B. 1977; 198:1–59. [PubMed: 20635]
16. Jones JP, Palmer LA. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J. Neurophysiol. 1987; 58:1187–1211. [PubMed: 3437330]
17. Kayama Y, Riso RR, Bartlett JR, Doty RW. Luxotonic responses of units in macaque striate cortex. J. Neurophysiol. 1979; 42:1495–1517. [PubMed: 115968]

18. Kinoshita M, Komatsu H. Neural representation of the luminance and brightness of a uniform surface in the macaque primary visual cortex. *J. Neurophysiol.* 2001; 86:2559–2570. [PubMed: 11698542]
19. Klein SA, Levi DM. Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. *Journal of the Optical Society of America.* 1985; 2:1170–1190. [PubMed: 4020514]
20. Kulikowski JJ, Bishop PO. Fourier analysis and spatial representation in the visual cortex. *Experientia.* 1981; 37:160–163. [PubMed: 7238742]
21. Lee TS. Image representation using 2D gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1996; 18:1–13.
22. Marcelja S. Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America.* 1980; 70:1297–1300. [PubMed: 7463179]
23. Marr D, Hildreth E. Theory of edge detection. *Proc. Roy. Soc. Lond., B.* 1980; 207:187–217. [PubMed: 6102765]
24. MacEvoy SP, Kim W, Paradiso MA. Integration of surface information in primary visual cortex. *Nat. Neurosci.* 1998; 1:616–620. [PubMed: 10196570]
25. Parker AJ, Hawken MJ. Two-dimensional spatial structure of receptive fields in monkey striate cortex. *Journal of the Optical Society of America.* 1988; 5:598–605. [PubMed: 3404310]
26. Peng X, Van Essen DC. Peaked encoding of relative luminance in macaque areas V1 and V2. *J. Neurophysiol.* 2005; 93:1620–1632. [PubMed: 15525807]
27. Petkov N. Biologically motivated computationally intensive approaches to image pattern recognition. *Future Generation Computer Systems.* 1995; 11:51–465.
28. Petkov N, Kruizinga P. Computational models of visual neurons specialized in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. *Biological Cybernetics.* 1997; 76:83–96. [PubMed: 9116079]
29. Polat U, Tyler CW. What pattern the eye sees best. *Vision Research.* 1999; 39:887–895. [PubMed: 10341942]
30. Potzsch M, Kruger N, von der Malsburg C. Improving object recognition by transforming gabor filter responses. *Network: Computation in Neural Systems.* 1996; 7:341–347.
31. Ringach DL. Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *J. Neurophysiol.* 2002; 88:455–463. [PubMed: 12091567]
32. Roe AW, Lu HD, Hung CP. Cortical processing of a brightness illusion. *Proc. Natl. Acad. Sci. USA.* 2005; 102:3869–3874. [PubMed: 15738406]
33. Rossi AF, Paradiso MA. Temporal limits of brightness induction and mechanisms of brightness perception. *Vision Research.* 1996; 36:1391–1398. [PubMed: 8762758]
34. Rossi AF, Paradiso MA. Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *J. Neurosci.* 1999; 19:6145–6156. [PubMed: 10407050]
35. Rossi AF, Rittenhouse CD, Paradiso MA. The representation of brightness in primary visual cortex. *Science.* 1996; 273:1104–1107. [PubMed: 8688096]
36. Victor, JD.; Knight, BW. Simultaneously band and space limited functions in two dimensions, and receptive fields of visual neurons. In: Kaplan, E.; Marsden, J.; Sreenivasan, KR., editors. *Springer Applied Mathematical Sciences Series.* Springer; 2003. p. 375-420.
37. Wachtler T, Sejnowski TJ, Albright TD. Representation of color stimuli in awake macaque primary visual cortex. *Neuron.* 2003; 37:681–691. [PubMed: 12597864]
38. Wallis G. Linear models of simple cells: Correspondence to real cell responses and space spanning properties. *Spatial Vision.* 2001; 14:237–260. [PubMed: 11817739]
39. Watson AB, Barlow HB, Robson JG. What does the eye see best? *Nature.* 1983; 302:419–422. [PubMed: 6835375]

Figure 1.

Simple balanced Gabor weight functions, $g(r) / g(0)$ vs. r , (left) and Hankel transforms, $H_g(\rho)$ vs. ρ , (right) for four exponents $\gamma_R = 0.75$ (red), 1.5 (green), 3.0 (blue), 6.0 (purple). As the exponent γ_R increases the weight $g(r)$ broadens and the Hankel transform $H_g(\rho)$ narrows. Note that the weight functions are strictly positive and monotonic, and are not oscillatory.

Figure 2.

Left panels compare the traditional cosine-type Gabor receptive field function (purple) to the corresponding simple balanced Gabor receptive field function (red) with identical gaussian components, that is, with identical weight functions (or envelopes, shown in gray). Right panels show corresponding Fourier transforms. All receptive field functions are cosine-type. The imbalance of the traditional Gabor is indicated by the nonzero values of its Fourier transform at the origin. Comparisons for $\gamma_R = 0.75, 1.5, 3.0, 6.0$ are illustrated. Note that the traditional and simple balanced Gabor receptive field functions become increasingly similar (i.e., the balancing parameter approaches 0) as γ_R increases (i.e., as effective spatial frequency bandwidth decreases), becoming virtually identical for $\gamma_R \gtrsim 3$.

Figure 3.

Left panels illustrate orientation max-response functions (in degrees) and right panels plot spatial frequency max-response functions for simple balanced Gabor receptive field functions as they vary with carrier spatial phase: 0 deg (cosine-type, red), 45 deg (mixed-type, green) and 90 deg (sine-type, blue). Response curves for $\gamma_R = 0.75, 1.5, 3.0, 6.0$ are illustrated. Note that as exponent γ_R increases max-response curves become independent of carrier spatial phase ϕ_R and that sine-type fields always give larger responses than cosine-type fields with corresponding parameters. The cosine-type spatial frequency max-response matches the Fourier transform of Figure 2 (differ only by a scale factor).

Figure 4.

Four pairs of plots illustrating nonsimple balanced Gabor weight functions, $g(r) / g(0)$ vs. r , (left) and Hankel transforms, $H_g(\rho)$ vs. ρ , (right) for four exponents $\gamma_R = 0.75, 1.5, 3.0, 6.0$ (with $\psi_R = 0$). Each plot shows three curves. The solid curves plot weight functions when c_R is set to a positive (blue) or negative (green) boundary value. These extreme curves (blue/green) are interchanged relative to the corresponding extreme curves for $\psi_R = \pi / 2$ (Figure 7). The red curve replots the weight function of the simple balanced Gabor ($c_R = 0$) for comparison. Note that nonsimple Gabor weight functions exhibit an oscillatory behavior and can take on negative values. The departure from the simple balanced Gabor weight function is most pronounced for large values of γ_R . As the exponent γ_R increases the weight $g(r)$ broadens and the Hankel transform $H_g(\rho)$ narrows.

Figure 5.

Four pairs of plots showing normalized balanced Gabor receptive field functions,

$\frac{\lambda_R^2}{2\pi g(0)} R(x_1, 0)$ vs. x_1 / λ_R (left), and associated weight functions, for four exponents, $\gamma_R = 0.75, 1.5, 3.0, 6.0$ (with $\psi_R = 0$), together with their Fourier transforms (right). All receptive field functions are cosine-type. Each plot shows three cases: $c_R = 0$ (simple balanced Gabor,

red), and $c_R = \frac{\pm \gamma_R^2}{\gamma_R^2 + 4\pi^2}$, representing upper (blue) and lower (green) boundary values of this coefficient. Weight functions (left panel) are in gray. Note that the amplitude of the receptive field functions is less than one for small values of γ_R due to the balance parameter. With respect to the Fourier transform $F_R(s_1, 0)$, note that: 1) as γ_R increases the transform narrows (consistent with decreasing bandwidth); 2) as γ_R increases the range of variation of the curves increases (consistent with the increased range of boundary values of c_R). This is consistent with an increased range of variation in spatial frequency bandwidth (within the overall narrowing). Note the clear variation in the Fourier transforms for small values of s_1 (corresponding to low spatial frequencies). Curiously there is no apparent corresponding variation in the receptive field functions themselves for small values of γ_R , indicating the importance of the large scale receptive field structure. Finally, note that because nonsimple Gabor weight functions exhibit an oscillatory behavior and can take on negative values, this can produce a phase-reversal of the receptive field function.

Figure 6.

Four pairs of plots showing orientation max-response functions in degrees (left) and spatial frequency max-response functions (right) for nonsimple balanced Gabor receptive field functions as they vary with c_R for four exponents, $\gamma_R = 0.75, 1.5, 3.0, 6.0$ (with $\psi_R = 0$),

where $c_R = 0$ (simple balanced Gabor, red) and $c_R = \frac{\pm \gamma_R^2}{\gamma_R^2 + 4\pi^2}$, corresponding to the max (blue) and min (green) boundary values of this coefficient. For each value of c_R there is plotted a triplet of curves corresponding to receptive field phase, $\phi_R = 0$ deg (cosine-type), 45 deg (mixed-type) and 90 deg (sine-type). Note that as γ_R increases the dependence of both orientation and spatial frequency response on receptive field phase ϕ_R , (holding c_R fixed) steadily decreases, becoming virtually independent of receptive field phase ϕ_R at $\gamma_R = 6.0$. For each value of γ_R , the cosine-type spatial frequency max-response function matches the Fourier transform of Fig. 5 (differ only by a scale factor). In particular, the plot for $\gamma_R = 6.0$ matches the corresponding plot for Fig. 5 because virtually no variation with carrier phase occurs here.

Figure 7.

Four pairs of plots illustrating nonsimple balanced Gabor weight functions, $g(r) / g(0)$ vs. r , (left) and Hankel transforms, $H_g(\rho)$ vs. ρ , (right) for four exponents $\gamma_R = 0.75, 1.5, 3.0, 6.0$ (with $\psi_R = \pi / 2$). Each plot shows three curves. The solid curves plot weight functions when c_R is set to a positive (blue) or negative (green) boundary value. These extreme curves (blue/green) are interchanged relative to the corresponding extreme curves for $\psi_R = 0$ (Figure 4). The red curve replots the weight function of the simple balanced Gabor ($c_R = 0$) for comparison. Note that nonsimple Gabor weight functions exhibit an oscillatory behavior and can take on negative values. The departure from the simple balanced Gabor weight function is most pronounced for large values of γ_R . As the exponent γ_R increases the weight $g(r)$ broadens and the Hankel transform $H_g(\rho)$ narrows.

Figure 8.

Four pairs of plots showing normalized balanced Gabor receptive field functions,

$\frac{\lambda_R^2}{2\pi g(0)} R(x_1, 0)$ vs. x_1 / λ_R (left), and associated weight functions, for four exponents, $\gamma_R = 0.75, 1.5, 3.0, 6.0$ (with $\psi_R = \pi/2$), together with their Fourier transforms (right). All receptive field functions are cosine-type. Each plot shows three cases: $c_R = 0$ (simple

balanced Gabor, red), and $c_R = \frac{\pm \gamma_R^2}{\gamma_R^2 + 4\pi^2}$, representing upper (blue) and lower (green) boundary values of this coefficient. Weight functions (left panel) are in gray. Note that the amplitude of the receptive field functions is less than one for small values of γ_R due to the balance parameter. With respect to the Fourier transform $F_R(s_1, 0)$, note that: 1) as γ_R increases the transform narrows (consistent with decreasing bandwidth); 2) as γ_R increases the range of variation of the curves increases (consistent with the increased range of boundary values of c_R). This is consistent with an increased range of variation in spatial frequency bandwidth (within the overall narrowing). Note that the variation in the Fourier transforms for small values of s_1 evident for nonsimple balanced Gabors with weight functions where $\psi_R = 0$ (Fig. 5) is here most prominent for values of s_1 slightly less than 1.0. The corresponding variation in the receptive field functions themselves is now most prominent in the secondary (rather than primary) lobes. Finally, note that because nonsimple Gabor weight functions exhibit an oscillatory behavior and can take on negative values, this can produce a phase-reversal of the receptive field function with a mild case occurring for $\gamma_R = 6.0$ (blue).

Figure 9.

Four pairs of plots showing orientation max-response functions in degrees (left) and spatial frequency max-response functions (right) for nonsimple balanced Gabor receptive field functions as they vary with c_R for four exponents, $\gamma_R = 0.75, 1.5, 3.0, 6.0$ (with $\psi_R = \pi/2$),

where $c_R = 0$ (simple balanced Gabor, red) and $c_R = \frac{\pm \gamma_R^2}{\gamma_R^2 + 4\pi^2}$, corresponding to the max (blue) and min (green) boundary values of this coefficient. For each value of c_R there is plotted a triplet of curves corresponding to receptive field phase, $\phi_R = 0$ deg (cosine-type), 45 deg (mixed-type) and 90 deg (sine-type). Note that as γ_R increases the dependence of both orientation and spatial frequency response on receptive field phase ϕ_R (holding c_R fixed) steadily decreases, becoming virtually independent of receptive field phase ϕ_R at $\gamma_R = 6.0$. For each value of γ_R , the cosine-type spatial frequency max-response function matches the Fourier transform of Fig. 7 (differ only by a scale factor). In particular, the plot for $\gamma_R = 6.0$ matches the corresponding plot for Fig. 7 because virtually no variation with carrier phase occurs here.

Figure 10.

Four pairs of plots illustrating bandlimited weight functions, $b(r) / b(0)$ vs. r , (left) and Hankel transforms, $H_b(\rho)$ vs. ρ , (right) for support parameter values $s_R = 1.0, 0.85, 0.7, 0.5$. Each plot shows three curves which plot weight functions when the order of the Bessel weight ν_R is set to 2.0 (green), 3.5 (red), and 5.0 (blue). Note that Bessel weight functions exhibit oscillatory behavior and take on negative values. As the support parameter s_R decreases, the weight $b(r)$ broadens and the Hankel transform $H_b(\rho)$ narrows.

Figure 11.

Four pairs of plots showing bandlimited Bessel receptive field functions, $\frac{\lambda_R^2}{2\pi b(0)} R(x_1, 0)$ vs. x_1 / λ_R (left), and their Fourier transforms, $F_R(s_1, 0)$ vs. $\log_{10}(\lambda_R s_1)$ (right), for support parameter values $s_R = 1.0, 0.85, 0.7, 0.5$. Each plot shows three curves which plot bandlimited weight functions when the order of the Bessel weight ν_R is set to 2.0 (green), 3.5 (red), and 5.0 (blue). Note that unlike balanced Gabor functions, the amplitude of the bandlimited receptive field function at the origin is equal to the weight function for all values of s_R because the balance parameter is 0. With respect to the Fourier transform $F_R(s_1, 0)$, note that: 1) unlike balanced Gabor receptive field functions, the maximum of the transform always occurs when $\lambda_p = \lambda_R$; 2) as support parameter s_R values decrease the receptive field function narrows in effective bandwidth; 3) for a particular support parameter value, as the order of the Bessel weight ν_R increases the receptive field function narrows in effective bandwidth; 4) like balanced Gabor receptive field functions, the oscillatory behavior of the bandlimited weight function can give rise to phase reversals of the receptive field function; an example is evident at $s_R = 0.7$.

Figure 12.

Four pairs of plots showing orientation max-response functions in degrees (left) and spatial frequency max-response functions (right) for bandlimited Bessel receptive field functions as they vary with $s_R = 1.0, 0.85, 0.7, 0.5$, and $\nu_R = 2.0$ (green), 3.5 (red), and 5.0 (blue). Unlike balanced Gabor receptive field functions, the max-response of bandlimited receptive field functions is independent of the receptive field phase ϕ_R . As support parameter σ_R values decrease the receptive field function narrows in effective bandwidth. For a particular support parameter value, as the order of the Bessel weight ν_R increases the receptive field function narrows in effective bandwidth.