Linear and nonlinear systems analysis of the visual system: Why does it seem so linear?∗
A review dedicated to the memory of Henk Spekreijse

Robert Shapley ∗

Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003, USA

A R T I C L E  I N F O

Article history:
Received 25 June 2008
Received in revised form 15 September 2008

Keywords:
Retina
Visual pathway
Contrast
Linearity
Visual cortex
System analysis

A B S T R A C T

Linear and nonlinear systems analysis are tools that can be used to study communication systems like the visual system. The first step of systems analysis often is to test whether or not the system is linear. Retinal pathways are surprisingly linear, and some neurons in the visual cortex also emulate linear sensory transducers. We conclude that the retinal linearity depends on specialized ribbon synapses while cortical linearity is the result of balanced excitatory and inhibitory synaptic interactions.

1. Introduction

Linear and nonlinear systems analysis are tools that can be used to study communication systems. While it may seem that these methods are more appropriate for systems engineered by humans, in fact powerful insights about the biological mechanisms of sensory systems – especially vision – have been obtained through the use of systems analysis. This review paper makes this point by examples drawn from experiments inspired by systems analysis applied to the vertebrate retina and to the mammalian primary visual cortex (V1). The experimental results together with systems analysis reveal that the retina has specialized cellular and molecular mechanisms that enable the retina to provide a faithful, linear transformation of the visual image. After examining the retinal transformation, we will consider how V1 cortex reconstructs and then selectively distorts the neural image.

Many sensory neurons can be understood as stimulus–response transducers that are driven by sensory stimulation from the outside world. Such a neuron is quiet or in a background state in the absence of stimulation. Then, when presented with an appropriate stimulus, the neuron is either activated above its background level of activity or in some cases suppressed below background in a more or less consistent manner from one stimulus presentation to the next. This consistency is called stationarity. When the stimulation ceases, the sensory neuron’s activity relaxes back to the background state. This property is called finite memory. A response that is stationary and of finite memory applies to most sub-cortical sensory neurons that have been studied and so we can consider them as stationary, finite-memory transducers. In the sensory areas of the cerebral cortex there are neurons that behave as sensory transducers according to how we have defined the term here, though not all cortical neurons fit this description. A cortical neuron that is involved in memory or decisions or the initiation of action will have some activity that is not stimulus driven, and therefore such a neuron will not fit neatly into the definition of a transducer neuron. Linear and nonlinear systems analysis techniques that we will be discussing in this paper are only applicable to neurons of the transducer type. Nevertheless, there are many neurons that can be understood as transducers and it is worth analyzing them in order to understand how neuronal networks can explain aspects of behavior.

Analysis of the visual system leads to the surprising conclusion that linearity is a rare and (apparently) prized commodity in neural signal processing. One reason that nonlinearity in neural information processing is the default is that neural communication is mainly through synaptic transmission, and most synapses are very nonlinear. The retina has very special synapses, the ribbon synapses that I will discuss later, and these specialized synapses appear to enable the retina to make use of linearity of signal...
transmission. The visual cortex seems to adopt a different approach because it does not have the luxury the retina has of handling continuous signals transmitted through ribbon synapses. The visual cortex must deal with the nonlinearity imposed by the spiking mechanisms of spiking neurons that feed it visual input. As I will describe, the cortex developed an intricate signal-balancing act to reconstitute a linear visual signal in the cortex. Cortical linearity is not simply the default result of convergence of excitatory inputs but rather requires extensive cortical computations. Thus the simplicity and elegance of linear systems are created in the visual system, and presumably also in other sensory pathways, by special efforts – by specialized synapses, or specially balanced networks. At the conclusion of this review paper I will offer some ideas about why the visual system works so hard to create, and then to reconstitute, a linearly filtered version of the visual world.

2. Linearity and nonlinearity in the vertebrate retina

Henk Spekreijse was a leader and an innovator in the application of systems analysis techniques to vision. He realized very early the importance of characterizing sensory transducers as linear or nonlinear. His early paper on linear and nonlinear analysis of visual responses in the goldfish retina (Spekreijse, 1969) applied an insightful method of linearizing neuronal responses with auxiliary signals to overcome the spike threshold nonlinearity of spiking neurons in the retina, the retinal ganglion cells. Fig. 1 from his 1969 paper summarizes many of his results on the spike-rate responses of goldfish ganglion cell to sinusoidal light modulation. But before we consider Spekreijse’s specific findings and their implications, we will discuss briefly why he used sinusoidal modulation of signals to study linearity and nonlinearity in retinal ganglion cells.

3. Linear systems and sinewaves

We can learn the principles of linear systems analysis from the analysis of the simplest linear transducers: linear, single-input, single-output, or LSISO systems. An LSISO system will respond to a brief pulse of input of unit area with a response $h(t)$, its impulse response. For an LSISO system, once we know the impulse response we know all there is to know about how the LSISO system will respond to any input. This is because linear systems obey the principle of superposition. Superposition means that the response of the cell to a sum of two stimuli, $x + y$, must equal the sum of the responses to the individual stimuli. Any realizable input can be decomposed into a sum of pulses at different times, of different heights. Based on superposition, the response of an LSISO system to this sum of pulses is simply the sum of its impulse responses to each of the pulses scaled appropriately. This process of summation is usually called convolution. Convolution is the basis for the linear synthesis of responses of LSISO transducers to any input (see Bracewell, 2000). This is why if we can measure $h(t)$, the impulse response, then the LSISO system is understood completely.

Sinewaves are favorite stimuli to use in linear systems because they pass through unchanged in waveform and frequency, and are simply scaled and phase-shifted. One can prove this using convolution (Bracewell, 2000). Therefore, if a system’s response includes sinusoidal components not in the input, this is a certain indicator of some kind of nonlinearity, and is given the name “distortion”. If a system responds to a sinusoidal input with an undistorted sine-wave output at the same frequency as the input, then it could be linear – at least such a system is emulating a linear system’s behavior to sinuosids. This is the reason Henk Spekreijse in 1969 was presenting sinewave-modulated light to the goldfish retina and monitoring the spiking rates of ganglion cells – to test for linearity or nonlinearity in retinal signal processing.

Fig. 1. Ganglion cell responses in the goldfish retina (from Spekreijse, 1969). The responses demonstrate the linearizing effect of an auxiliary signal (first column), internal noise (second column) and spontaneous spike discharge (third column). The calibration bars are 20 spikes/bin. The bin duration was 625 μs. (A–G) The lowest points represent the 0 spikes/s. (H) For the lowest points represent a firing rate of the order of 4 spikes/s. All responses are from red “OFF” ganglion cells.
What Spekreijse (1969) reported was that usually there was a lot of distortion in the average spike-rate response of goldfish retinal ganglion cells (for example Fig. 1B), but that under just the right conditions one could coax the retina to respond in a linear manner, that is to produce an undistorted sinewave output response (Fig. 1C, F, and H) to sinewave light modulation (Fig. 1A). The nonlinearity he was exploring was the spike threshold of the ganglion cells, and the way he made the ganglion cells to yield linear responses was by using very small signals (Fig. 1F) or by finding ganglion cells that had unusually high spontaneous spike-firing rates (Fig. 1H) or by using an auxiliary signal added to the sine-wave input as in Fig. 1B. Hughes and Maffei (1966) had also found sinusoidal spike-rate responses in cat retinal ganglion cells. In the cat retina, the ganglion cells usually have moderately high spontaneous spike-firing rate, like the goldfish case in Fig. 1H, and sinusoidal modulation of the firing rate was observable without the need for auxiliary stimuli.

4. Mechanisms of retinal linearity—ribbon synapses

There are many different implications of the ground breaking work of the Spekreijse (1969) paper, but for the purposes of this review paper I want to draw attention to the fact that it showed that the vertebrate retina could respond like a linear system to visual inputs all the way from the photoreceptors through to the retinal ganglion cells. Many vision scientists myself included have taken this result for granted and focused on the many interesting features of the nonlinear stages of signal processing in the retina, but the linearity exhibited in Fig. 1 is most important for vision. Understanding what cellular and molecular mechanisms are needed for such linear signal processing remains a challenge for neuroscience.

What the retina is usually facing is a visual scene that has a mean light level and modulations above and below this steady level. In other words there is a continual bombardment of the retina by photons from the environment, and modulations of the rate of the photon flux comprise the visual message. In order to achieve a linear retinal response to modulations of the photon flux, the photoreceptors must be linear transducers around an operating point, and this they manage to do (Baylor, Hodgkin, & Lamb, 1974; Tranchina, Sneyd, & Cadenas, 1991).

The synapses from photoreceptors to bipolar cells and from bipolar cells to retinal ganglion cells must also operate in a linear manner around the mean level of synaptic release determined by the mean level of photon flux. I propose that the ribbon synapse depicted in Fig. 2 (copied from DeVries, Li, & Saszik, 2006) is the crucial organelle that the retina uses to achieve linear signal processing (picking up a suggestion from Hochstein & Shapley, 1976a). The ribbon synapse is specialized for continual release of substantial quantities of glutamate packaged in synaptic vesicles (reviews in Heidelberger, 2007; Heidelberger, Thoreson, & Witkovsky, 2005; Witkovsky, Thoreson, & Tranchina, 2001). In the case of cone photoreceptors and their synapses, the light level at which cones usually operate is high enough that each cone will receive many photons/s and the mean level of synaptic transmitter release will be determined by the mean light level. Then modulations of light intensity around this operating point, as in natural scenes or in the laboratory experiments that mimic natural conditions, will modulate the amount of synaptic transmission above and below the mean level. In rod photoreceptor synapses the situation is different because the operating point is usually quite different from that of cones: lower mean light levels that usually provide less than 1 photon/s to the rod. Then synaptic release is dominated by the dark current that depletes the rod (Hagins & Yoshikami, 1975), and photon events cause a large transient reduction in synaptic release. At the usual operating point of the rod synapse, nonlineairties in synaptic transmission do not matter because the retina is acting more as a photon counter, summing synaptic events coming from many different rods. Indeed, the rod–bipolar synapse is suspected of having a threshold nonlinearity (Sampath & Rieke, 2004) but the threshold may be low enough that even the rod synapse is approximately linear (Heidelberger et al., 2005; Robson, Maeda, Saszik, & Frishman, 2004; Witkovsky et al., 2004). But for the normal operating range of the rod the amount of nonlinearity of the rod–bipolar synapse does not much matter because the retina can operate in a linear manner by counting photon evoked events. So for the remainder of our discussion let us consider the cone synapses and how they manage to transmit an undistorted signal at higher photon arrival rates.

The photoreceptor ribbon synapse is presynaptic to both bipolar and horizontal cells. Evidence for linearity of synaptic transmission by means of sinusoidal driving of the photoreceptors by sinewave modulation of illumination has come mainly from experiments on horizontal cells because of the greater difficulties recording from bipolar cells though there are some data from bipolar cells that support the idea of the linearity of synaptic transmission from cones to bipolar cells (Marmarelis & Naka, 1973; Toyoda, 1974).

Spekreijse and Norton (1970) presented some of the earliest evidence for the linearity of the photoreceptor ribbon synapses in their study of horizontal cell responses in the goldfish retina. They used different temporal waveforms of illumination and observed horizontal cell response waveforms (Fig. 3). For our purposes we will focus on the sinusoidal waveforms in Fig. 3A. It is evident that the goldfish horizontal cells produced sinusoidal responses to sine-wave modulation of the light, with little obvious distortion.

Later, Tranchina, Gordon, Shapley, and Toyoda (1981) performed another experiment using sinusoidally modulated light to study horizontal cells in the turtle retina (Fig. 4). In Fig. 4, the left panel shows the response waveform to a moderately high contrast (or modulation depth) of 0.5. As a test of how linear the transduction of the cone-horizontal cell synapse was, Tranchina et al. calculated the Fourier amplitude spectrum of the horizontal cell response – or in other words the best fitting sinewave at the modulation frequency (1st harmonic), at 2× the modulation frequency (2nd harmonic), at 3× the modulation frequency (3rd harmonic), and so on. The reason why they used the harmonics to test for nonlinearity is as follows. We can write an expression for the light stimulus in these experiments

\[ I(t) = I_0 + I_1 \cos \omega t. \]
The amount of harmonic distortion is therefore one quantitative measure of amount of nonlinearity, and this measure was used by Tranchina et al. (1981) to assess cone–horizontal cell transmission. The relative heights of the harmonics in the turtle horizontal cell response are depicted in the middle graph of Fig. 4. The 1st harmonic is tall and all the other harmonics are very short – consistent with a very linear transduction across the synapse. In other words, the cone ribbon synapse has very little harmonic distortion. In the right panel of Fig. 4 the amplitudes of the 1st harmonic and 2nd harmonic responses are plotted vs stimulus contrast of the sinusoidal modulation. The 1st harmonic is proportional to contrast while the 2nd harmonic is negligible up to the highest contrast measured (around 0.5). In other experiments Tranchina et al. observed absence of harmonic distortion up to contrasts as high as 0.8.

There is some difference of opinion, possibly caused by species differences, in the precise mechanism of linear signal transduction at the cone ribbon synapse. Investigators of the goldfish retina and of the catfish retina have found evidence for the idea that horizontal–cone feedback modulates synaptic transmission at the ribbon synapse, and this modulation makes synaptic transmission more linear (Kraaij, Spekreijse, & Kamermans, 2000; Sakai & Naka, 1987; Verweij, Kamermans, & Spekreijse, 1996). Kraaij et al. (2000) hypothesized that the early response of horizontal cells to a step increment of illumination was driven “open-loop” by the cones, while the later part of the step response was “closed-loop”, that is, influenced by horizontal–cone feedback. When they studied open-loop and closed-loop responses to different light intensities, they found that the ratio of horizontal/cone responses (what they called the gain of the cone synapse) was fairly linear – that is, proportional to light intensity – for the closed-loop response (Fig. 5B), but quite nonlinear for the open-loop condition (Fig. 5A).

Sakai and Naka (1987) found more evidence for horizontal–cone feedback’s influence on the linearity of synaptic transmission when they studied the size dependence of horizontal cell responses to noise-modulated light (Fig. 6). They were recording the membrane potential of horizontal cells in the catfish retina, and stimulating the retina with a light stimulus that was a constant plus a Gaussian white noise (GWN) signal. Such a light stimulus appears to be flickering randomly in time. The use of such random signals to probe transducer properties is a powerful tool for studying linear and nonlinear systems. Schellart and Spekreijse (1972) reported one of the first experiments that used noise and cross-correlation in their study of goldfish retinal ganglion cells. Sakai and Naka (1987) were using the Wiener kernel approach that had been pioneered by Marmarelis and Naka (1973). They calculated the first-order Wiener kernel by cross-correlation of the noisy input signal with the noisy neuronal response. The first-order
Then Sakai and Naka made a test of how linear the cone–horizontal cell transmission was, by comparing the response of the best-fitting linear system with the response of the neuron. They did this comparison by calculating the power spectrum of the horizontal cell’s response, and comparing it with the predicted power spectrum of the response of the best-fitting linear system to the same GWN stimulus, as shown in Fig. 6. In the figure the data are plotted with a solid line and the predicted response of the linear system is the dashed line. There were two experimental conditions: S means only a centrally located spot was stimulated by light and the rest of the visual field was dark; S/A means the spot was modulated in the presence of steady illumination of the surrounding region by an annulus of light that had the same mean retinal illumination as the spot. The data in Fig. 6 show that in the S/A condition, the power in the horizontal cell’s response would be explained completely by the best-fitting linear system implying that the horizontal cell’s response was linear and also, therefore, cone–horizontal cell synaptic transmission was linear. But in the S condition when only the spot was illuminated, there was a significant discrepancy between the linear prediction and the measured response, implying significant nonlinearity in cone–horizontal cell transmission when only the spot was illuminated. It is known that horizontal cells have very large visual receptive fields and so it is reasonable to suppose that there was more sustained horizontal cell response in the S/A than in the S condition. Then the difference between the S and S/A conditions in linearity of cone–horizontal cell transmission could have been caused by the different amounts of horizontal cell response leading to different amounts of horizontal–cone feedback. A similar deviation between the linear prediction and measured power spectrum was reported for catfish bipolar cells (Sakai & Naka, 1987).

The generality of the role of horizontal–cone feedback in linearity of synaptic transmission is called into question by the same kind of spot-annulus experiment applied to the turtle retina (ChapPELL, Sakuranaga, & Naka, 1985), as shown in Fig. 7. As in the catfish retina, the presence of annular illumination affects the waveform of the measured first-order kernel (Fig. 7A), and also the power spectrum of the horizontal cell’s response (Fig. 7B). But in the turtle horizontal cell, the comparison of predicted power spectra with measured spectra (Fig. 7B) yields a qualitatively different result: for the S condition as for the S/A condition, the linear prediction matches the measured power spectrum. This implies that the entire response of the turtle horizontal cell was consistent with linear transduction, whether the stimulus was a spot that did not engage much horizontal–cone feedback, or whether it was a modulated spot plus steady annulus that did evoke substantial feedback. Further work will be needed to clarify why there is a species difference between fish and reptiles (and other vertebrates) in the mechanisms of linearity at the cone ribbon synapses.

One concern is that OFF-center (flat) bipolar cells that make synapses at the base of the cone pedicle might have different kinds of synaptic transmission from the ON-center bipolar cells that make invaginating synapses closely apposed to the synaptic ribbons, as diagrammed in Fig. 2 (from DeVries et al., 2006). Even if the ribbon synapses conferred linearity on the ON-pathway, one might wonder whether the OFF-pathway could have the same property of linearity if the synapses were not ribbon synapses. The results of DeVries et al. (2006) put such concerns to rest. Investigating cone–bipolar transmission in the retina of the ground-squirrel, they found that both invaginating and flat bipolar cells are driven by transmitter released by ribbon synapses. The synaptic current is slightly delayed in the flat OFF bipolar cells compared to invaginating ON bipolar cells but this is almost compensated by the fact that metabotropic responses in invaginating ON bipolar cells are slightly delayed compared to the ionotropic responses in OFF cells. The single source of synaptic input to ON and OFF bipolar

![Fig. 5. Cone–horizontal transmission in goldfish retina (Kraaij et al., 2000). The "open-loop" early response (A) and "closed-loop" from the later response (B) are gain-characteristic functions from cone to horizontal cells (HC) (solid lines). The open-loop gain-characteristic is highly nonlinear, whereas the closed-loop gain-characteristic is nearly linear.](image1)

![Fig. 6. Horizontal cell responses in the catfish retina (Sakai & Naka, 1987). Power spectra of the (spot) light stimuli, of the spot-evoked responses (continuous lines), and of the linear models (dashed lines). Spectra marked by "S" were by the spot alone and those marked "S/A" were by the same spot of light in the presence of a steady annular illumination. The mean square errors were 45% and 10%, respectively, for the models for a spot alone and for the spot in the presence of a steady annular illumination.](image2)

Wiener kernel is the impulse response of the linear system whose response to the GWN signal is the best fit to the neuron’s response.
cells solves the problem of how there could be corresponding degrees of linearity of signal processing in the ON and OFF-pathways.

Another finding about bipolar cells worth mentioning in the context of linearity of signal processing is the quasi-linearity of the membrane potential of isolated salamander bipolar cells driven by injected currents (Mao, MacLeish, & Victor, 1998). While the bipolar cells have voltage gated currents that make them nonlinear transducers, around an operating point, a mean voltage level, their responses are quite linear in waveform and amplitude. The membrane nonlinearity reported by Mao et al. (1998) may be involved in contrast adaptation or light adaptation – that is gain control processes that set the operating point around which the neurons may be modulated in a linear manner.

5. Retinal linearity–retinal ganglion cells

The linear transduction of the visual image would go to waste if signals were not transmitted without distortion to the retinal ganglion cells, the outputs of the retina. Now we will review the linearity of signal processing in cat retinal ganglion cells. The evidence suggests that linearity is preserved along some pathways all the way through the retina. A way to study this question is to examine the linearity of spatial summation of neural signals. A sinusoidal grating pattern, examples of which are shown in Fig. 8, is a useful tool for this task. If a ganglion cell is simply adding up neural signals, and there is no difference in the time course of the response from the different signal sources, then positions can be found at which introduction and withdrawal of the grating produce no response (left column, Fig. 8; Enroth-Cugell & Robson, 1966). At these null positions the grating is placed so that introduction of the pattern produces as much net positive signal from one side of the cell’s receptive field as it produces net negative signal from the other side of the field; the two signals of equal magnitude but opposite sign cancel when added. The X cell data in Fig. 8 are from an OFF-center X cell; they are average spike-rate responses as functions of time, at two peak and two null spatial phases. Null positions can be found for X retinal ganglion cells in the cat retina, but not for Y cells (right side of Fig. 8) that show signs of nonlinear signal distortion.

The early experiments of Enroth-Cugell and Robson (1966) employed a temporal modulation signal that was a step of contrast from zero to a fixed value. The time period when the contrast was stepped to a nonzero value is indicated in Fig. 8 as the horizontal line. But one can use other temporal modulation signals, and in particular Hochstein and Shapley (1976a) suggested the use of sinusoidal temporal modulation in order to test the linearity of the retina by using the kind of temporal waveform analysis we have been considering in this review article. Thus, if one chooses the spatio-temporal stimulus to be a contrast-reversal grating,

$$I(x, t) = I_0 + I_1 \sin(\phi x + \zeta) \cos(\omega t)$$

where \( \phi = 2\pi k \) where \( k \) is spatial frequency in cycles/deg; spatial phase \( \zeta = 2\pi Q \) where \( Q \) is spatial offset in a fraction of a spatial cycle; temporal modulation frequency \( \omega = 2\pi f \) where the units of \( f \) are cycles/s or Hz. One can examine the average spike-rate modulation as a function of time and test for harmonic distortion just as investigators studied harmonic distortion in horizontal and bipolar cells. In response to a sinusoidal contrast reversal stimulus, the responses of X retinal ganglion cells follow a sinusoidal function of spatial phase as illustrated in Fig. 9 (from Enroth-Cugell & Robson, 1984). The data are from an experiment on a cat OFF-center X retinal ganglion cell. The sinusoidal spatial phase (position) dependence is a consequence of linearity of spatial summation (Hochstein & Shapley, 1976a). But in keeping with our focus on temporal waveform, it is quite salient that the temporal waveform of the ganglion cell’s average spike-rate response was a sinewave at the modulation frequency of the stimulus, and second harmonic distortion was very small (Hochstein & Shapley, 1976a).

The experimental outcome is very different in cat Y cells as shown in Fig. 10 (Enroth-Cugell & Robson, 1984), where the second harmonic component of the response is as large as the 1st. The relative sizes of the 1st and 2nd harmonics in the responses of Y cells depend on visual stimulus parameters like spatial frequency \( k \), contrast \( I_1/I_0 \), and mean illumination \( I_0 \).

Fig. 11 (Hochstein & Shapley, 1976a) illustrates the spatial frequency dependence of a cat Y cell’s 2nd/1st harmonic ratio, as well as the spatial-phase invariance of the 2nd harmonic response that is also evident in Fig. 10. Much scientific effort went into analyzing the nonlinear retinal pathways that drive Y cells in order to explain among other things the spatial frequency dependence, and the spatial-phase invariance, of the 2nd harmonic responses (Hochstein & Shapley, 1976a; 1976b; Victor & Shapley, 1979a; 1979b, among others) but these do not concern us in this paper. It is important to note that ON-center and OFF-center X cells are comparably linear in their response waveforms as indicated by harmonic analyses. The X cell in Fig. 9 was an OFF-center cell, while the X cell that provided the data for Fig. 11A was ON-center. One piece of explanation of this is, as we wrote above, that both ON-center and OFF-center bipolar cells are receiving input from cones through ribbon synapses (DeVries et al., 2006). Also, the results
imply that bipolar–ganglion cell signal processing is equivalently linear in both ON- and OFF-center pathways.

Similar analyses of retinal linearity revealed the existence of retinal ganglion cells that resembled cat X cells in the retinas of many other vertebrate species: eel (Shapley & Gordon, 1978); rabbit (Caldwell & Daw, 1978); frog (Gordon & Shapley, 1978); goldfish (Levine and Shechner, 1979; Bilotta & Abramov, 1989); mudpuppy (Tuttle & Scott, 1979). It is reasonable to conclude that in the vertebrate retina there usually is a linear pathway for visual signals all the way from photoreceptors to retinal ganglion cells. I suggest that the ribbon synapses in photoreceptors are one component, and the ribbon synapses between bipolar and ganglion cells (Dowling & Boycott, 1966) are another necessary component of this linear pathway.
Sum of sinusoids analysis of linearity in retinal ganglion cells

Jonathan Victor developed a different approach for studying linear and nonlinear systems, and applied it first to studying the retina (Victor & Knight, 1979; Victor & Shapley, 1980; Victor et al., 1977). This is the use of a sum of sinusoids (SOS) as a temporal modulation signal. For instance, instead of studying X cells with a contrast-reversal grating where the spatio-temporal stimulus was

\[ I(x, t) = I_0 + I_1 \sin(\phi x + \xi) \cos(\omega t) \]

Victor used as a stimulus

\[ I(x, t) = I_0 + I_1 \sin(\phi x + \xi) \sum_j \cos(\omega_j t) \]

where the index \( j = 1, 2, \ldots, 8 \), and \( \omega_j = 2\pi f_j \) where the units of the \( f_j \) are cycles/s or Hz. The SOS stimulus is illustrated in Fig. 12. The eight different sinusoids in the sum are drawn there, as well as the sum at the bottom. The frequencies \( f_j \) were selected very carefully. In fact, usually Victor et al., 1977 used frequency sets where \( f_j = (2^{j-1} - 1)/T \) (where \( T \) was the period of stimulation, approx. 30 s) so that sums and differences of pairs of frequencies were all distinct. Distinct output frequencies were useful for the following reason. In a linear system, the response to SOS must be a weighted sum of the input frequencies \( f_j \). Therefore, SOS with distinct output frequencies is useful for systems analysis because, as illustrated in Fig. 13, if one Fourier analyzes the response of a system to SOS, the linear part of the response will appear at the input frequencies \( f_j \) but nonlinear response components will appear at the sum frequencies \( f_j + f_k \) or at the difference frequencies \( f_j - f_k \). The sum and difference frequencies are intermodulation distortions caused by the same nonlinearities that cause harmonic distortions; indeed second harmonic distortion \( 2f_j = f_j + f_j \) is a special case of intermodulation distortion. In a way, the use of SOS is an extension of Spekreijse’s concept of linearizing (Spekreijse, 1969), but with many auxiliary signals. Fig. 14 illustrates how we constructed the temporal frequency response of the best fitting first-order system (what we called the first-order frequency kernel, Victor & Shapley, 1979a) from the neuronal response to SOS modulation of a spatial pattern.

The use of SOS allowed us to test for linearity of signal processing definitively. The experimental result is presented in Fig. 15 (Victor et al., 1977). Data from one cat X cell are displayed in the right hand column in the figure, and Y cell data are on the left. The first-order responses are the curves in the upper panels, and the second-order intermodulation responses are graphed as contour lines depicting equal height along surfaces in the lower panels. For the purposes of this paper I will focus on the X cell’s data. While we were able to measure a large first-order response (10 spikes/s amplitude of response) for the X cell, there was negligible second-order response at any pairwise intermodulation frequency, either among the sum frequencies or among the difference frequencies. This contrasts with the Y cell data where there was a large hill of response in the sum frequency quadrant, and a smaller but still measurable hill of response in the difference...
frequency quadrant. These results were typical; across the X cell population we studied, first-order responses were 5–10\% bigger in summed amplitude than second-order responses (Victor & Shapley, 1979a).

Before completing the part of this paper devoted to retinal signal processing, I need to add qualifications. The X cell pathway is remarkably linear for a pathway in a neural network but the retina is by no means a linear network. What we have considered is the linearity of signal transmission around an operating point of average illumination and average contrast. But the retina is also equipped with adaptation mechanisms to adjust input–output transductions. The first-order response of X cells is affected in a nonlinear way by mean illumination (reviewed in Shapley & Enroth-Cugell, 1984) and by mean contrast (Benardete & Kaplan, 1999; Shapley & Victor, 1978; Sakai, Wang, & Naka, 1995; Victor, 1987). Under normal operating conditions in viewing real scenes, the retina is working around an operating point, and the retina’s linearity around its operating point is remarkable.

7. Linear signal processing in simple cells of the primary visual cortex (V1)

Neurons in the primary visual cortex are classified as simple or complex, depending on how they respond to visual stimuli. If the response of the cell depends on the stimulus in an approximately linear fashion, the cell is termed “simple” and if nonlinear then it is called a “complex cell”. For instance, in response to visual stimulation by the temporal modulation of grating patterns, the linearity of simple cells includes: (1) a sensitive dependence on the spatial phase of the grating as in X retinal ganglion cells, (2) very little 2nd harmonic distortion. The responses of complex cells are very different: (1) they are spatial-phase-insensitive, and (2) their responses are predominantly 2nd harmonic (DeValois, Albrecht, & Thorell, 1982; Spitzer & Hochstein, 1985; Fig. 16). The linear dependence on visual stimuli of the simple cell might be assumed to be a simple consequence of convergence of excitatory drive from lateral geniculate nucleus (LGN) cells (Hubel & Wiesel, 1962; Reid & Alonso, 1995; Fig. 17). However, such a feedforward model fails because of the nonlinearities of the LGN cells.

The situation of the visual cortex is not like that of the retinal ganglion cells that are only two ribbon synapses away from the visual image. Spiking neurons, the ganglion cells and the LGN cells, are in the pathway to the cortex. Rectification caused by the spike-firing threshold produces nonlinear distortion of LGN re-
sponses for stimulus contrast >0.2 (Shapley, 1994; Tolhurst & Dean, 1990). Therefore, it is an open and important question, how can there be simple cells in the visual cortex? Wieliaard, Shelley, Mclaughlin, and Shapley (2001) offered an answer to this question by studying a large-scale neuronal network model of layer 4C\alpha\ in macaque primary visual cortex, V1. The choice of lateral connectivity within the Wieliaard model is motivated not by Hebbian-based ideas of activity-driven correlations (Troyer, Krukowski, Priebe, & Miller, 1998), but by an interpretation of the anatomical and physiological evidence concerning cortical architecture. The crucial distinguishing features of the model, derived from biological data, are that the local lateral connectivity is nonspecific and isotropic, and that lateral monosynaptic inhibition acts at shorter length scales than excitation (Callaway, 1998; Callaway & Wiser, 1996; Fitzpatrick, Lund, & Blasdel, 1985; Lund, 1987). We have tried to make a distinction between orientation preference and orientation selectivity. In the model, orientation preference is conferred on cortical cells by the convergence of output from many LGN cells (Reid & Alonso, 1995), with that preference laid out in pinwheel patterns (Blasdel, 1992a, 1992b; Bonhoeffer & Grinvald, 1991; Maldonado, Godecke, Gray, & Bonhoeffer, 1997). Mclaughlin, Shapley, Shelley, and Wieliaard (2000) showed that the orientation selectivity of cells in such a model of 4C\alpha\ is greatly enhanced by lateral cortico-cortical interactions.

Wieliaard et al. (2001) found that neurons in the large-scale network model behaved like V1 simple cells and this was a result of the cancellation of nonlinear LGN excitation by cortico-cortical inhibition. The neurons in the model were integrate-and-fire neurons with synaptic excitatory and inhibitory conductances modeled after those in real cortical cells. Therefore, the model outputs included intracellular conductances and membrane potential as a function of time. Its results could be compared with intracellular recordings of the membrane potentials of visual cortical cells, for instance the results of Jagadeesh, Wheat, Kontsevich, Tyler, and Ferster (1997); Fig. 18. The intracellular results are impressive because they reveal that even the intracellularly recorded membrane potential contains very little 2nd harmonic distortion (Fig. 18) even though such distortion is present in the LGN input (Fig. 19) and the Wieliaard et al. (2001) model can account for the intracellular results too. Please note that Fig. 18 from the original paper contains two cycles of temporal modulation, so the response is at the fundamental frequency, not the 2nd harmonic.

The nonlinearity in the model is caused by LGN rectification, initially. The spatial arrangement of LGN cell receptive field centers that provide input to each cortical cell is as segregated ON–OFF subregions as in the feedforward model (Reid & Alonso, 1995). The ON–OFF segregation confers an orientation preference on the input to each cortical cell, and this preference (in the model) is laid out in pinwheel patterns. Additionally, the center of the receptive field of each cortical cell (created through the aggregate LGN input) is randomized. This was done to account for diversity in the location of this receptive field center, and possible random variations in the spatial symmetry of the ON–OFF subregions. The spatial arrangement of the LGN input confers a preferred spatial phase on the LGN input of each cortical cell.

From cortical cell to cell this spatial-phase preference is distributed randomly over a broad range, as has been found in experimental measurements (DeAngelis, Ghose, Ozhawa, & Freeman, 1999).

In response to contrast reversal, the summed LGN drive in the model has (for 100% contrast modulation) the generic spatial phase and time dependence shown in Fig. 19. Notice that for each phase, the sinusoidal shape is significantly distorted. The absolute maxima of response occur at either 1/4 cycle or 3/4 temporal cycle of modulation, and the orthogonal phase case (with the lowest peak heights of response) has two peaks per cycle, that is, frequency-doubling or 2nd harmonic response. The 2nd harmonic is a consequence of the rectification in LGN firing rate. It is this 2nd harmonic that has to be eliminated for V1 cells to be simple cells.

The way the 2nd harmonic in the excitatory LGN drive is eliminated in simple cells in the Wieliaard et al. (2001) model is by strong cortico-cortical inhibition. Thus, simple-cell-like intracellular and spike-rate responses to contrast reversal stimuli occur in the model because the model’s cortico-cortical inhibitory conductances have significant 2nd harmonic modulations that cancel the 2nd harmonic coming from the input. This statement is borne out by examining the inhibitory conductances in model neurons, as shown in Fig. 20, where the 2nd harmonic in the conductance waveform is obvious. But why are such 2nd harmonic modulations present in the model’s inhibitory conductance? The reason is that the jth model neuron receives spikes from many other cortical neurons, each of which is responding individually in a manner sensitive to the spatial phase of its own LGN drive. This individual spatial phase dependence arises because each of these cortical neurons is driven by LGN excitation, and each summed LGN drive will have its own temporal waveform that will be one of those sketched in Fig. 19. The excitation of each LGN cell is maximal at 1/4 or 3/4 temporal cycle. Because the cortico-cortical input to the jth neuron is an average over many such spatial-phase-sensitive responses, some of which peak at 1/4, some at 3/4 cycle, this results in a total...
cortico-cortical conductance that peaks at 1/4 and 3/4 of the temporal cycle, and consequently has significant 2nd harmonic content. In summary, the cortico-cortical conductances have large, phase-insensitive, 2nd harmonic modulations because the isotropic cortical architecture of the model allows an averaging over the activity of many cortical neurons, and thus, indirectly averages over the many preferred spatial phases of the LGN input [as suggested by the results in DeAngelis et al. (1999)], which peak at 1/4 and 3/4 cycle. This “phase averaging” by the network is similar to that used in a model for complex cells (Chance, Nelson, & Abbott, 1999). It should be emphasized that although we have invoked phase averaging as the mechanism for producing

Fig. 16. Simple cell responses to grating contrast reversal from DeValois et al. (1982) (with author’s permission). (A) Macaque monkey simple cell, spike-rate response to contrast reversal of a sine grating at 2 Hz modulation. Position of the grating in the visual field is specified in degrees of spatial phase: one spatial cycle of the grating pattern is 360°. (B) Macaque complex cell response to the same contrast reversal stimulus. The response amplitude shows little variation with spatial phase, and there are two response peaks per cycle of temporal modulation—this is the 2nd harmonic component.

Fig. 17. Classic feedforward model from LGN to simple cells in V1 cortex. Adapted with permission from Hubel and Wiesel (1962). Four LGN cells are drawn as converging onto a single V1 cell. The circular LGN receptive fields aligned in a row on the left side of the diagram make the receptive field of the cortical cell elongated.
frequency-doubled cortico-cortical inhibitory conductance, this state of cortical activity arises from the dynamics of the system in a way consistent with its architecture.

The result of the interplay between excitation and inhibition in the model is that model cortical cells, unlike their LGN input, behave like simple cells in the contrast reversal experiment. Fig. 21a–c shows data from an excitatory model neuron located near a pinwheel center. In Fig. 21a both the “in-phase” and “orthogonal phase” membrane potential responses are shown. Here, the spike and reset mechanism of this neuron has been turned off—blocked—so that the waveform of stimulus-modulated membrane potential, $V_b$, can be seen more easily and compared with the experimental data of Jagadeesh et al. (1997) where spikes were filtered out. Thus, the averaged waveforms of the membrane potentials in Fig. 21a should be compared with those shown in Fig. 18. There is a good degree of similarity. Extracellular spike counts for this same model neuron (spike and reset now on) are displayed in Fig. 21b and c, as cycle-averaged histograms, and these are comparable to the simple cell data in Fig. 16. Fig. 21a-c show that model neurons have the linearity seen experimentally in simple cells. The response at the peak spatial phase is predominantly at the 1st harmonic of temporal modulation. The spike rate is not modulated at the second harmonic when the stimulus is at the “null phase”, and the membrane potential shows very little 2nd harmonic component in its null phase response, consistent with experimental measurements. If this explanation of simple cell function is correct, it means that the cortex is emulating some of the behavior of a linear system without actually being a network of linear elements. The apparent linearity of response time course and spatial summation are the result of a balance between nonlinear excitation and nonlinear inhibition, according to this way of thinking.
To study the effect of lateral cortico-cortical interactions in the model, we shut them off. Fig. 21d–f shows the results of a simulation with all network interactions shut off but with the LGN input the same as in the full network simulation shown in Fig. 21a–c. For the null phase condition, notice the large amplitude of the 2nd harmonic in both the spike-rate response and the membrane potential. This 2nd harmonic response is inherited from the LGN input (as seen in Fig. 19 in the null phase LGN responses). The responses of an uncoupled model neuron are much larger than seen in the living cortex, because of the removal of strong inhibition in the model. Another approach to cortical modeling is to choose different input and internal noise parameters for the uncoupled model neurons to fit the background and peak firing rates of the real cortex. We did this and investigated the responses of what we called a "feedforward" neuron with much weaker LGN drive than in the full model. The results of the simulation for the feedforward neuron are shown in Fig. 21g–i. Compared with both the responses of the feedforward and uncoupled neurons, the membrane potential of the fully coupled neuron has a much smaller 2nd harmonic component, because of cortico-cortical interactions.

Our view that simple cells must be created by network interactions, and are not simply the default result of excitatory convergence, is supported by many experiments that reveal that the "simple" property of simple cells, the linearity, can be affected by unbalancing excitation and inhibition in the cortical network. Cells could be shifted from simple to complex by say weakening inhibition. This was observed (Fregnac & Shulz, 1999; Murthy & Humphrey, 1999). The data of Murthy and Humphrey (1999) are particularly relevant. They stimulated their cat simple cells with grating contrast reversal as in our modeling and observed marked frequency doubling in spike rates of simple cells when bicuculline (which weakens GABA-ergic inhibition) was infused. The opposite effect, namely transferring a cell from the complex to simple group by increasing inhibition, was reported by Bardy, Huang, Wang, FitzGibbon, and Dreher (2006).

Simple and complex cells were first discovered in cat visual cortex (Hubel & Wiesel, 1962), and their existence confirmed subsequently in macaque V1 (DeValois et al., 1982; Hubel & Wiesel, 1968). Simple cells have been found in the primary visual cortex of many other species of mammals: owl monkeys (O’Keefe et al., 1998), baboons (Kennedy, Martin, & Whitteridge, 1985), tree shrews (Kauffman & Somjen, 1979), rats (Burne et al., 1984; Girman et al., 1999), mice (Draeger, 1975), rabbits (Glanzman, 1983), and sheep (Kennedy, Martin, & Whitteridge, 1983). Where there is a visual cortex, there are simple cells.

The modeling work suggests that the linear behavior of simple cells arises as a consequence of network activity in the cortical network. Why is linearity the cortical network’s goal? One idea is that, for visual perception, cortical cells must resolve and represent key spatial properties such as surface brightness and color, and also the perceptual organization of a scene. The existence of simple cells that respond selectively to spatial phase and monotonically to signed contrast are a requirement for the representation of surface properties. The large corpus of work on spatial vision requires lin-
ear spatial mechanisms to explain how patterns are detected separately and in combination (Graham, 1989; Wandell, 1995). Moreover, theories of color vision implicitly assume the existence of simple cells whenever they postulate the necessity of numerical computations of (signed) edge contrast (Wandell, 1995). A different function of vision also requires cells like simple cells. Scene organization requires computation of depth order that in turn depends on computation of stereoscopic depth and also of pictorial occlusion. Both stereo (Anzai, Ohzawa, & Freeman, 1999a, 1999b) and occlusion (Anderson, 1997) computations appear to require cortical representation of signed edge contrast. Also, the perception of salient contours embedded in a noisy field of distractors has been shown to be sensitive to spatial phase and thus contrast sign of the elements of the contour (Field, Hayes, & Hess, 2000). Such neural computations would seem to require the linearity that only simple cells provide. Visual perception needs simple cells for important, basic functions. The retina and the visual cortex strive to create linear spatio-temporal elements to perform those functions. Recent work by Dr. Patrick Williams in my laboratory (Williams & Shapley, 2007), suggested that V1 reconstructs linear-look simple cells in the input layer 4C, and these cells respond in an approximately linear manner not only to sinusoidal temporal modulation but also to steps of contrast of the kind they might experience after eye movements. The work of Friedman, Zhou, and von der Heydt (2003) and also Williams and Shapley (2007) shows that V1 proceeds to make nonlinear but contrast-polarity-sensitive neurons in the upper layers of V1. The ultimate design goal appears to be neuronal sensitivity to contrast polarity. Further investigations in my laboratory by Drs. Chun-I Yeh and Dajun Xing are revealing more about the nonlinear transformation from layer 4C to layer 2/3 but the results are still preliminary. But it appears from what V1 cortex does and from the linearly filtered signals sent from retinal ganglion cells that a stage of neuronal processing that emulates a linear spatio-temporal filter is useful for further visual image processing in the visual system. Usually the determination of the linearity or nonlinearity of signal processing is a first step in linear and nonlinear systems analysis. Then the goal of systems analysis is to obtain a set of measurements that identify or characterize the system structure by comparing the filtering properties of a system (linear or nonlinear) with a model system (for instance, see Enroth-Cugell & Robson, 1966; Schellart & Spekreijse, 1972; Spekreijse, 1969). What I have tried to show here is that even the first step of systems analysis may be very revealing about the function of a neural system and may lead us to uncovering unsuspected functional constraints in the neuronal mechanisms of the retina and of the brain.

Acknowledgments

This work was supported by Grants R01 EY01472 from the US National Eye Institute, and NSF 0745253 from the US National Science Foundation. Thanks also to Drs. Shaul Hochstein, Daniel Tranchina and Jim Wieland for helpful comments on earlier versions of the paper.

References
