Linearity and non-linearity in cortical receptive fields

Robert M. Shapley

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

Abstract. Visual neurons in striate (V1) cortex have been studied as feature detectors or as spatiotemporal filters. A useful way to distinguish between these two conceptual approaches is by studying the way in which visual signals are pooled across space and time. Many neurons in layer IV of striate cortex exhibit linear spatial summation and their response time course is consistent with linear temporal summation. Neurons in supragranular and infragranular layers sum signals in a non-linear manner. A particularly important non-linearity seen in many cortical complex cells is non-linear summation along an axis parallel to their preferred orientation. This leads to responsiveness to 'illusory contours', borders defined by texture differences only. These and other results on non-linear summation of chromatic and achromatic signals imply that V1 cortex performs sophisticated and complex image processing and is not simply an array of spatiotemporal filters.

1994 Higher-order processing in the visual system. Wiley, Chichester (Ciba Foundation Symposium 184) p 71–87

The visual system needs to perform sophisticated computations on neural signals evoked by visual images in order to enable us to recognize objects and navigate through the world. Determination of something seemingly as simple as the colour of reflecting objects under different conditions of illumination is quite difficult. Other non-trivial visual tasks include determination of the direction of motion of moving objects and the shape of an object, moving or still. Flat pictures and even line-drawings of solid objects are remarkably recognizable, too; this ability needs to be explained.

Many lines of evidence indicate that the visual areas of the cerebral cortex are involved in the neural computations that support visual performance. The primary visual area—V1, or striate cortex—is the initial visual cortical processing area and has been studied extensively for a long time. Yet, we have only the beginning of an understanding of what signal processing is actually going on in V1. In this paper, I shall discuss what we know about how visual signals are combined by neurons in V1 to make up their receptive fields. The nature
of signal summation reveals quite a bit about the function of visual neurons and about the function of the cortical neural network.

Consideration of visual receptive fields is linked inextricably with the issue of the role of primary visual cortex in visual information processing. The original ideas of, for instance, Hubel & Wiesel (1962) about feature detection were not framed in a quantitative manner, although the appeal of feature detection as a cortical function endures. The concept of the visual cortex as a bank of narrow-band spatial filters has been predominant as a rationale for research in this field (reviewed by Shapley & Lennie 1985). However, there are many problems with the idea of the cortex as a Fourier analyser. One is that it does not solve the problems of object segmentation and identification, which have to be deferred to ‘higher areas’.

There has been evidence for some time that the visual cortex does something more interesting than filtering; evidence presented below about illusory contour responsiveness fits in with this new concept of the visual cortex as an image processor rather than as a filter bank. This new concept is also supported by recent psychophysical and perceptual work by, among others, Nakayama & Shimojo (1990), who suggest that sophisticated image processing about the nature of surfaces must be happening at a fairly low level in the visual system. Recent work by Lamme et al (1992, 1993) also supports this idea. The case for viewing the primary visual cortex as an image processing engine rather than as a passive spatiotemporal filter bank follows.

Linearity and cortical simple cells

Before dealing with the complexities of non-linear image processing, one needs to explain responses of ‘simple cells’ in striate cortex. Simple cells are the orientation-selective neurons in V1 that appear to sum visual signals in a linear or approximately linear, manner (Hubel & Wiesel 1962, Movshon et al 1978a, Spitzer & Hochstein 1985a, De Valois et al 1982). I will discuss my own results mainly on simple cells in macaque monkey (Macaca fascicularis) striate cortex but much of the classical work was done on simple cells in cat area 17 (primary visual cortex in cat). Recent analysis indicates that complex neural network are required to account for linear response properties of simple cells. Simple models based on linear summation of excitatory lateral geniculate nucleus (LGN) inputs (e.g. Hubel & Wiesel 1962) were not designed to and do not account for the degree of linearity observed in both cat and monkey simple cortical cell (e.g. Movshon et al 1978a, De Valois et al 1982, Reid et al 1991).

The problem with a simple ‘summation-of-excitation’ model of simple cell is that the cells’ linearity of spatial summation, as assessed by a modified null test with sinusoidal gratings undergoing contrast reversal (Enroth-Cugell & Robson 1966, Hochstein & Shapley 1976a, Movshon et al 1978a, Reid et al 1991) persists up to quite high values of stimulus contrast. The summation-of-excitation
model would predict substantially greater non-linearity than is observed in simple cells. This indicates that the cortical network is connected in a sophisticated manner to achieve linearity in simple cells by means of precise balancing of excitation and inhibition.

Tolhurst & Dean (1990) pointed out that responses of cat LGN neurons are often non-linear at contrasts of 0.2 (20%) and higher, because of high contrast sensitivity and only a modest maintained firing rate—leading to rectification within the LGN because the modulated spike rate is clipped at zero impulses per second. Thus, a simple additive model for cat cortical simple cells would be expected to show non-linear distortion, such as frequency doubling of responses to contrast reversal of gratings, whereas many observations on cat simple cells indicate apparent linearity of spatial summation without distortion (Movshon et al 1978a, Spitzer & Hochstein 1985a, Reid et al 1991). Michael Hawken, David Grosos and I have now done similar experiments in monkey, reported below, that agree with the results on cat cortex.

(he basic experiment to test linearity of spatial summation is measurement of neural responses to contrast reversal of a sine grating as a function of the spatial phase of the grating. The stimulus arrangement for such an experiment is shown in Fig. 1. The process of grating contrast reversal is illustrated by showing the stimulus spatial waveform at four moments in time, during one cycle of contrast reversal. Two spatial phases are shown—‘peak’ spatial phase, which should elicit the maximal response from the hypothetical visual neuron for which the sensitivity profile is drawn beneath the stimulus, and ‘null’ spatial phase, at which stimulus modulation elicits no response. This is the procedure that Shaul Hochstein and I developed for studying X and Y cells in the LGN and retina (Hochstein & Shapley 1976a), based on the earlier null test of Enroth-Cugell & Robson (1966). For a linear visual neuron for which spatial and temporal responses are separable (that is, a single temporal response function characterizes all positions in the receptive field), the amplitude of response to contrast reversal of a sine grating should be a sinusoidal function of spatial phase. Figure 2 illustrates such linear behaviour for two parvocellular LGN neurons in macaque monkey (Kaplan & Shapley 1982).

The pattern of responses to the stimulus of Fig. 1 is quite different for visual neurons that receive visual inputs which have gone through a non-linear transduction. There are two characteristic features: distortion of the response waveform by second harmonic distortion and absence of a ‘null’ response (Enroth-Cugell & Robson 1966, Hochstein & Shapley 1976a). It is important to note one particular variant of this non-linear behaviour, namely non-linear responses that have the same amplitude independent of spatial phase. This is not always the case, but has been observed in cat Y retinal ganglion cells (Hochstein & Shapley 1976a,b, Victor & Shapley 1979) and in some cat complex cells (Movshon et al 1978b, Spitzer & Hochstein 1985b). Under the assumption that all the subunits are spatially similar, this spatial phase invariance is
FIG. 1. Grating contrast reversal. The curves at the top of the figure represent the one-dimensional luminance profile of the grating at four instants in time. The luminance profile is a sinusoid. Its amplitude is modulated in time by a slow modulation signal which in these experiments is a sinusoidal function of time. The solid curve is the luminance profile of the grating at the crest of the temporal modulation signal. The coarsely dashed curve is the luminance profile at the trough of the temporal modulation signal. The finely dashed curves are stimulus spatial profiles at intermediate values of temporal modulation. The solid curve below the stimulus represents a sensitivity profile of a visual neuron. The vertical line is the midpoint of the receptive field. In the stimulus situation on the left, labelled ‘0° spatial phase’, the crest of the spatial sinusoid is lined up with the midpoint of the receptive field; this is the condition of maximal sensitivity for a linear neuron. In the situation depicted on the right, denoted ‘90° spatial phase’, a zero-crossing of the grating is lined up with the midpoint of the sensitivity profile; the null position for a linear neuron.


When we run a grating contrast reversal experiment on monkey V1 cortical cells, we obtain results characteristic of linear spatial summation in about half the cells, including simple cells in all layers of V1 (Fig. 3). There is a clear, strong dependence of response amplitude on spatial phase, and it is approximately sinusoidal. There is significant harmonic distortion, as indicated in the inset by the square symbols, but this is explained entirely by the clipping of the response at zero spikes/second because of the low maintained firing rate of cortical cells. These results are particularly significant in simple cells from layer IVb that are presumed to receive their visual input predominantly from neurons in layer IVcα which in turn are mainly driven by magnocellular neurons. This is
FIG. 2. Spatial phase responsiveness of parvocellular X-like neurons in the macaque monkey lateral geniculate nucleus. Fundamental Fourier responses to grating contrast reversal are plotted as a function of spatial phase of the grating pattern. The stimulus contrast was 0.35; temporal frequency, 4 Hz; spatial frequency, 2 cycles/degree. The filled and empty circles represent data from two different parvocellular neurons. The solid curve in the lower panel is the best-fitting sinusoid. Response phases are plotted above; straight lines were drawn through the phase data. From Kaplan & Shapley (1982) with permission.

because the experiment was done with a (peak) contrast of 0.6. Figure 4 shows that the response of a typical magnocellular neuron becomes non-linear at much lower contrasts (0.08 for this cell); above a contrast of 0.6 the response is highly clipped and consequently distorted. How could responses like that at 0.6 contrast in Fig. 4 be summed to produce the quasi-linear behaviour in Fig. 3? Tolhurst & Dean (1990) proposed an explanation for visual neurons in cat area 17 in terms of the push–pull model (Hubel & Wiesel 1959, Palmer & Davis 1981). In this model, overlapping areas of ON-excitation and OFF-inhibition are pooled with an offset region of overlapping ON-inhibition and OFF-excitation. Pooling of ON and OFF signals with opposite sign restores linearity from clipped LGN inputs in ON and OFF pathways. The explanation advanced for cat cortical neurons is also required to account for simple cells driven by magnocellular signals in macaque monkey V1 cortex. The role of inhibition is crucial in restoring linearity at the cortical level. Furthermore, the strengths of excitation and inhibition must be matched precisely by connection strength within the cortical network.
FIG. 3. Spatial phase dependence of the response of V1 cortical simple cell to grating contrast reversal. This was a neuron recorded from layer IVb. Grating contrast was 0.6; spatial frequency, 1.05 cycles/degree; temporal frequency, 8.45 Hz. Response histograms are shown; the numbers at the right of each histogram are the spatial phases in $\pi$ radians. In the inset on the left, fundamental and second harmonic amplitudes are plotted vs. spatial phase; circles are the fundamental Fourier amplitudes of the response while squares are the second harmonic amplitudes. The inset on the right shows fundamental and second harmonic responses plotted as vectors in polar coordinates to show amplitude and phase of the responses at each of the different spatial phases.

**Complex cells and illusory contours**

Many cortical cells in V1 do not exhibit the ‘linear’ pattern of responses seen, for example, in Fig. 3. These other neurons are all cast into the category of complex cells. There are many varieties of complex cells with characteristic differences within and between cortical layers (Hawken et al 1988). In monkey complex cells one sees some of the types observed in cat cortex (reviewed in Spitzer & Hochstein 1988), but also additional types, perhaps because of the complications of pooling of signals from parvocellular and magnocellular pathways (De Valois et al 1982, Lennie et al 1989). These many types of cells may be understood, at least qualitatively, as variants of a neuron that pools inputs from multiple parallel sources, in each of which there is a non-linear transduction. The variation in complex cells comes from the number of non-linear inputs (called subunits). The number of subunits is usually inferred from the pattern of responses of a complex cell in a grating contrast reversal experiment like Fig. 1.

D. Grosof, M. Hawken and I have recently reported that macaque complex cells can respond to subjective or illusory contours, such as borders defined by
the termination of a sine grating or the border between two out-of-phase gratings (Grosos et al. 1993). This was unexpected and illustrates the possible functional importance of the non-linearities in the V1 neural network. Results from an experiment on a V1 complex cell in layer V are shown in Fig. 5. In this case, the stimulus was a drifting grating pattern, but drifting in a directional parallel to the bars of the grating. The grating was bounded by blank regions at the
same mean luminance as the grating. When the boundary of the grating crossed the neuron's receptive field, the neuron was excited transiently. The neuron could be simply responding to the local luminance change at a single bar ending. To test this, we varied the spatial phase of the grating pattern with respect to the neuron's receptive field. As in the case of Figs 1-3, if the neuron were responding to local bar endings alone, there should be strong spatial phase dependence of response amplitude and waveform. Figure 5 shows that there is approximate spatial phase invariance. This means that complex cells of this type are integrating signals from along the contour to signal the appearance of a texture boundary. About half the complex cells we studied with this

![Graph](image)

**FIG. 5.** Responses of a complex cell in macaque VI to illusory contours. Shown are the average response histograms of the responses of a layer 5, directionally selective, complex cell to a terminated grating pattern (1.3 cycles/degree grating, 80% contrast), drifted at 6 degrees/second in the optimal edge direction; the spatial phase (in π radians) is written alongside each histogram. Each pattern drifted to the edge of the screen, wrapped around and returned to its starting position every 947 ms. A histogram of the spikes using a period of 947 ms shows that the cell responded to each of the two illusory contours passing over the receptive field in a period. The inset plot shows the amplitude of the first (F1, circles) and second (F2, squares) harmonics.
paradigm in macaque V1 responded vigorously to illusory contours of this type with spatial phase-invariant response amplitudes.

There are other kinds of illusory figures, like the one in Fig. 6, that we have not yet tested on cortical neurons in V1. Figures like this are interesting because of the gap that is bridged by the illusory contour.

Conclusions

The more closely one investigates the function of cortical V1, the less it looks like a simple bank of spatial filters and the more one has to respect the sophistication and 'logic' of neural information processing. I have selected only two salient examples from a wealth of new evidence that indicates a sophisticated image processing role for macaque V1. By the usual assumptions one may extend these results to human visual cortex. There are two main examples. (1) Linear information processing is achieved in simple cells by an elaborate summation and balancing of excitation and inhibition. This implies that it is important to the visual system to retain a linear information pathway in the cortex. A possible reason for this is the importance of linearity for calculating brightness and colour for spatial patterns. (2) Illusory contour responsiveness is achieved in V1 complex

FIG. 6. Illusory contours with gaps between inducers. This is a figure that is similar to illusory shapes invented by G. Kanizsa (cf. Gerbino & Kanizsa 1987). In this case, there is a gap between each pair of inducers that equals in length the contour defined by line terminations within each inducer.
cells by some sort of alignment of non-linear subunits, or cooperative interactions between subunits, in a way that has not yet been investigated. The importance of responsiveness to contours has been written about many times previously (see, for example, Petry & Meyer 1987).

In the study of how the visual cortex contributes to higher-order visual processing, further investigation of linear and non-linear signal summation in cortical neurons will almost certainly provide new insights.

Acknowledgements

Thanks to Michael Hawken and David Grosof for their participation in experiments reported here. I also thank Shaul Hochstein, Clay Reid and Nava Rubin for helpful discussions. My research has been supported by the US National Eye Institute, National Science Foundation, US–Israel Binational Science Foundation and by a MacArthur Fellowship.

References

Petry S, Meyer G (eds) 1987 The perception of illusory contours. Springer-Verlag, New York
Shapley RM, Lennie P 1985 Spatial frequency analysis in the visual system. Annu Rev Neurosci 8:547–583
Tolhurst DJ, Dean AF 1990 The effects of contrast on the linearity of spatial summation of simple cells in the cat's striate cortex. Exp Brain Res 79:582–588
Victor JD, Shapley RM 1979 The nonlinear pathway of Y ganglion cells in the cat retina. J Gen Physiol 74:671–689

DISCUSSION

Gilbert: I'm a little uneasy about your suggestion that your stimuli represent illusory contours for the reason that there is a contrast difference along the gratings you used. Although the average luminance difference between one side and the other integrated along the full pattern may be zero, you could account for the response of the cells to the pattern by the presence of subunits within a receptive field. There is an important distinction between the pattern and the kind of stimulus that von der Heydt & Peterhans (1989) used in studying illusory contours in V2, where the cell responds to contours that lie entirely outside the receptive field.

Shapley: The spatial phase experiment was designed to show that the neurons are responding to some extended property of the contour, as opposed to a local property.

Barlow: When two sinusoidal gratings of differing phase are placed side by side, or when there is just one grating that is terminated abruptly, complicated Fourier components are generated at orientations different from that of the original grating. How can you be sure that your units are not responding to these? This may not seem a serious problem if the period of the gratings is much smaller than the size of the receptive field or its subunits, but when the period is comparable or larger, there will be obvious edges at right angles to the grating
orientation. I do not see how you can be sure that your cells or their subunits are not responding to these.

Shapley: The grating spatial frequencies were approximately the same for studying the contour properties as for studying the normal orthogonal sensitivities in the preferred orientation. So these spatial frequencies were roughly of the order of magnitude of the central summation region.

Gilbert: Do you see any difference in kind between that sort of stimulus and the kind of illusory contour that you presented in Fig. 6?

Shapley: I do see a difference between different sets of illusory contours. I think the kind of illusory contour we used is an interesting stimulus to study because we have control over the spatial properties of the pattern in a way that I was trying to illustrate in the study of the linear receptive field properties. One could use this paradigm to study whether there were multiple subregions that were summing over space or whether a small discrete region was providing the major input to the cell.

Gilbert: In terms of the underlying mechanism, one might be tempted to think of different mechanisms for something that activates a cell yet is entirely outside the classical receptive field and the activation of subunits within the receptive field that sum in a non-linear fashion.

Shapley: We haven’t tested these cells to be sure that gap stimuli wouldn’t activate them. We have shown that if you do use stimuli of the kind that von der Heydt & Peterhans used, with rectangular wave gradients where there are gaps in contrast between the bars, you can also get robust frequency-doubled responses of the kind I described (Grososf et al 1993).

Gilbert: Are those gaps entirely bracketing a receptive field?

Shapley: Again, they are like the stimuli that von der Heydt & Peterhans used to obtain responses in monkey cortex and they are stimuli of the kind that G. Kanizsa, J. Kennedy, C. Ware, W. Ehrenstein and many other people have used for studying illusory contours (Petry & Meyer 1988). I agree that they are not identical to Kanizsa’s triangle, in the sense of having long blank regions, but we have to proceed in stages and try to do one experiment at a time.

Movshon: Dr. Gilbert, it’s also important to know that von der Heydt & Peterhans (1989) never actually show evidence of the kind that you’re claiming that Bob Shapley doesn’t show, which is that their stimuli are in fact outside the receptive field. What they do is show you a drawing of where they assert the receptive field was and as you (Gilbert 1994, this volume) told us very clearly, the size of the receptive field is dependent on how you measure it and under what conditions you test it.

Wilson: It would be interesting to see whether you can get similar contour responses with a significantly higher spatial frequency than the frequency of an orthogonal grating. Have you tried this?

Shapley: We haven’t been able to do that. We have studied these things as a function of spatial frequency along the contour and the resolution in that
dimension is not high, so it's not higher than in the orthogonal dimension. It might be interesting to do, but if you don't find similar responses, I am not sure that this tells you much except that the subregions that you are summing over are fairly coarse.

_Derrington:_ Have you looked at how the cell sums the length of the contour?

_Shapley:_ No, that is a very important thing to do, which is what Charles Gilbert is getting at—if you separate the inducers, over how long a gap will the response jump?

_Derrington:_ Or, how many bars do you need in your Vernier-offset gratings to elicit a good response?

_Shapley:_ Yes, I agree.

_Morgan:_ Do you need both sets of bars and thus a phase boundary? Cells with circularly symmetrical receptive fields will pick out the bar terminations because they respond more to terminations than along the length of the line. An oriented collector unit that summed the responses of these units would have the properties of the cells you describe. The cell would not need the phase boundary but would respond to either half of the stimulus. Are you saying that is what happens?

_Shapley:_ Yes, you do get a response from one half of the stimulus. A model of the kind that you suggest, where there is an alignment of subregions along the direction parallel to the contour, is a reasonable model for some kind of contour responses. I don't think that our data or anyone else's necessarily rule out such a model. In order to get the dynamics of the response that we see, you need to have these neural subregions not just randomly scattered, but rather tightly aligned.

_Morgan:_ I meant scatter along the axis of alignment to account for the phases.

_Shapley:_ I understand. But you would need to have either some way of registering the responses of these spatial subregions, some method of temporal synchronization, or some other sophisticated method in order to generate the synchronicity of the responses with the border. That in itself would be a level of sophistication of complex cell receptive fields that people haven't imagined before.

_Derrington:_ How sophisticated it is depends on how much summing you do along the length of the contour, of course.

_Movshon:_ Another way to phrase the question, in terms of the basic mechanisms that we think we understand in complex cells, is the following. Complex cells will respond to the introduction of a sinusoidal grating at any phase. So, one could view the half-stimulus that Michael Morgan referred to as the introduction along its length of one of these stimuli. Is there any difference between the character of the cell's responses to the introduction of the grating by being slid along the receptive field and to it being simply turned on in place, without contour? To what degree is contour itself critical and to what degree is merely the introduction of a contrast signal to which the cell responds critical?
Shapley: I think the answer lies in the synchronization of the response to the border passage.

Movshon: But that’s the arrival of the grating in either case—it arrives on the field. Imagine a photodiode placed in the middle of the receptive field. It would see a luminance change whenever the border crosses, as it would if we were to time-modulate the contrast.

Shapley: I don’t think the photodiode model is a bad one, but you would need to have multiple photodiodes arranged along the axis parallel to the optimal orientation of the pattern in order for it to have responsiveness to these contours.

Movshon: You don’t know that that’s true. The photodiodes have to be dispersed at right angles to the preferred orientation of the cell, because otherwise you wouldn’t get the response at all phases.

Shapley: But they don’t have to be aligned.

Movshon: I don’t believe they have to be aligned to your responses either—that’s the point.

Shapley: If they weren’t aligned, you would see a smeared out response in time.

Movshon: Right, but is the response more smeared out in time than you would predict from the length profile of the receptive field, or less?

Shapley: The stimulus being aligned doesn’t impose alignment of the subregions, because it’s a moving stimulus and so the response it evokes could be smeared out over time. The idea that there might be coincidence detection is a good one. The problem there is that you can also measure responses as a function of contrast and the contrast dependence does not show the highly accelerating non-linearity a coincidence model would require.

Lennie: I was surprised that you emphasized the linearity of operation so much. If we have learned anything substantial in the last few years about the operations of V1 it is that the non-linearities are much more pervasive than we thought. It’s clear, following particularly David Heeger’s analysis, and the contrast normalization work of Bonds (1991) and Geisler & Albrecht (1992), that the contrast response functions of simple cells are highly non-linear. Aren’t you overstating things substantially to put the linear operations anywhere near on a par with non-linear operations?

Shapley: Let me justify putting them on a par. I thought it was interesting that, given the inputs as they are, we should see as much linearity as we do. Any time you see something that’s against your expectations, you should try to explain it. I agree there are pervasive non-linear operations going on in the cortex, but it seemed to me that this was a case where it might be interesting to draw your attention to the fact that simple cells are linear and they have this characteristic of superposition. This is something that people have taken for granted but it requires specific neural interactions as a base.

Lennie: One of the reasons we didn’t notice these non-linearities before is that many of us, including myself, were drawing straight lines on plots of
response against contrast. It is not clear to me that, over the range where most
people have studied the things that you are calling linear behaviour, it’s really
linear at all.

Shapley: I said it emulated a linear mechanism. You see, for instance, the
elipses from our work—of the kind that Tony Movshon, Dave Tolhurst and
Ian Thompson found initially (Movshon et al 1978)—they certainly look like
they are emulating a linear operation. I didn’t choose the most linear ellipses.

Bergen: There are different aspects of linearity that may have very different
functional significance. In particular, one can separate (to some extent) the
spatial characteristics of linearity from its intensive characteristics. One test
traditionally used to diagnose linearity in visually driven cells is to look for a
position in the cell’s receptive field at which the introduction of an edge stimulus
elicits no response. A linear cell with inhibitory as well as excitatory regions
in its receptive field must necessarily have such a null position. The existence
of this null position, however, does not imply linearity. A truly linear cell will
also double its response when the contrast of the stimulus is doubled. The point
is that while a linear cell will have both of these properties, each one can exist
in a non-linear cell independent of the other. The spatial property can best be
described as phase or position sensitivity. The intensive property is simply
homogeneity of degree one.

The functional significance of each of these two properties is quite different.
If a cell is position sensitive, then the level of the cell’s response carries some
information about the position of the stimulus within the cell’s receptive field.
Homogeneity of the cell’s response is not important for this function. In fact,
in order to avoid confusing position variation with contrast variation, the cell
must (over some range) have a contrast-invariant response which is inconsistent
with homogeneity of degree one and therefore with linearity.

I think that Bob Shapley’s point is a very good one. I would paraphrase it
as saying that the spatial aspect of linearity that may be very important for
subsequent processing is something that doesn’t just happen. The visual system
must construct it rather carefully.

Barlow: Do you have a take home message about the significance of linearity
and non-linearity?

Shapley: The significance of linearity lies in the preservation of signature of
contrast. The usefulness of being able to know whether you are on the bright
side or dark side of the boundary is probably pretty important: you want to
go to some length to be able to retain that information, even to the point of
balancing ON and OFF inputs and going through the sort of somersaults that
the visual system seems to be going through in order to do this. It might be
particularly important for colour to be able to retain signature, so you know
that you are working with a dark red as opposed to a bright red and so on.
Signature of colour signals would be important for doing spatial computations
of colour in some sort of retinex-like later computation of spatial influences.
As far as non-linearity goes, its significance is for the purposes of determining the segmentation of the visual scene into objects. To know what’s in front and what’s behind, you don’t need to worry about which one’s brighter or darker; you really want to know whether there is a contrast gradient and, therefore, whether there is something in front, whatever its colour or brightness is. For that, these non-linearities blind to contrast signature are important for segmenting or breaking apart one thing from another. Both in chromatic and achromatic systems, even-order, non-linear stages that are just going for magnitude of difference are very important for figure/ground—saying something is different from something else. You start from that and then tack on the sign-sensitive signals that you preserved in these parallel linear pathways.

**Wilson:** But sign-sensitive is much more general than linearity. Are you just saying that the system needs both even-order and odd-order non-linearities?

**Shapley:** I think, too, that it’s more than just odd-order non-linearity, because if you want to do colour computations you need sign and magnitude.

**Wilson:** If you are thinking of raising some signal to an exponent that’s an integer, even integers produce a full-wave rectification of some sort and odd integers produce a first and third quadrant odd symmetrical function that preserves input sign.

**Shapley:** The leading term of any odd-order, non-linear expansion is a linear term, of course. If you want to do some undistorted faithful rendition across space of signals, you want to have mostly fundamental and less third-order input. On the evidence, the neurons look like they are, in some way, emulating a linear summation.

**Movshon:** There is a good deal of evidence that the visual system can, in fact, use a linearly filtered representation of signals. There are many cases where if you have an early strong non-linearity, even an odd-order one, there would then be interaction terms that arose that would have a perceptual consequence later on. Experiments like Newsome and I did with plaids (Movshon et al 1986) and experiments that people have done with compound and chequerboard gratings of various kinds (De Valois et al 1979), would all show much more strongly non-linear effects than they would do if there were a very important component of higher than order one. So, I think Bob Shapley’s case for linearity is more than just a case for contrast sign, it’s a case for decent approximation of linearity over a decent range of contrast. You need that signal—the point is that that signal alone isn’t going to solve a lot of interesting problems for you.

**Gilbert:** In contrast to the cat, where the input layer is composed entirely of simple cells, which project up to a layer composed of complex cells (lending plausibility to the idea that the subunits of complex cells are simple cells), in the monkey, it’s less clear. I’m not sure if it’s ever been demonstrated where the simple cells are in the monkey, and how they stand in relationship to the input and to the complex cells. Where do you find them?
Shapley: In the monkey, we see cells that are classically referred to as simple cells in all layers; they aren’t confined to one layer. We even see them in layer IVcβ—that is, oriented cells, predominantly responding at the fundamental frequency for a drifting grating. However, to get these beautiful sinusoidal dependencies on spatial phase, you don’t always observe such simple behaviour. The best examples I have are really these layer IVb and IVcα cells that are in the magnocellular pathway. We have some interesting examples in IVcβ, for instance, which are simple-like according to classical criteria, but in which spatial phase sensitivity shows substantial non-linearity. The most linear examples I have found so far have been in layer IVb.

Gilbert: I have not looked systematically, but that’s where I’ve seen them.

Shapley: We have also seen some nice examples of simple cells in layer VI of the monkey cortex (R. Shapley, M. J. Hawken & D. H. Grosos, unpublished results).

References