THE EFFECT OF CONTRAST ON THE TRANSFER PROPERTIES OF CAT RETINAL GANGLION CELLS

BY R. M. SHAPLEY AND J. D. VICTOR

From the The Rockefeller University, New York, New York 10021, U.S.A.

(Received 17 January 1978)

SUMMARY

1. Variation in stimulus contrast produces a marked effect on the dynamics of the cat retina. This contrast effect was investigated by measurement of the responses of X and Y ganglion cells. The stimuli were sine gratings or rectangular spots modulated by a temporal signal which was a sum of sinusoids. Fourier analysis of the neural response to such a stimulus allowed us to calculate first order and second order frequency kernels.

2. The first order frequency kernel of both X and Y ganglion cells became more sharply tuned at higher contrasts. The peak amplitude also shifted to higher temporal frequency at higher contrasts. Responses to low frequencies of modulation (< 1 Hz) grew less than proportionally with contrast. However, response amplitudes at higher modulation frequencies (> 4 Hz) scaled approximately proportionally with contrast. Also, there was a marked phase advance in these latter components as contrast increased.

3. The contrast effect was significantly larger for Y cells than for X cells.

4. The first order frequency kernel was measured with single sine waves as well as with the sum of sinusoids as a modulation signal. The transfer function measured in this way was much less affected by increases in contrast. This implied that stimulus energy at one temporal frequency could affect the response amplitude and phase shift at another temporal frequency.

5. Direct proof was found that modulation at one frequency modifies the response at other frequencies. This was demonstrated by perturbation experiments in which the modulation stimulus was the sum of one strong perturbing sinusoid and seven weak test sinusoids.

6. The shape of the graph of the amplitude of the first order frequency kernel vs. temporal frequency did not depend on the amplitudes of the first order components, but rather on local retinal contrast. This was shown in an experiment with a sine grating placed at different positions in the visual field. The shape of the first order kernel did not vary with spatial phase, while the magnitudes of the first order responses varied greatly with spatial phase.

7. Models for the contrast gain control mechanism are considered in the Discussion.
INTRODUCTION

The responses of retinal ganglion cells inform the brain of what the eye has seen. However, this neural message is not a simple transduction of the optical image. The image on the retina is transformed by spatial interactions between retinal interneurones, and this transformation is reflected in the discharge pattern of ganglion cells. Therefore, to comprehend the purpose and the functional machinery of the retinal network one needs to understand the activity of ganglion cells.

In our research we have concentrated on the responses of cat retinal ganglion cells to particular visual stimuli. These stimuli were spatial sine grating patterns which were amplitude-modulated by a temporal modulation signal which was a sum of sinusoids. We used sine gratings as spatial stimuli because they allowed us to dissect apart retinal mechanisms on the basis of spatial resolution (cf. arguments in Enroth-Cugell & Robson, 1966; Hochstein & Shapley, 1976a). We used a sum of sinusoids as a temporal modulation signal because it allowed us to tease apart components of the neural response which are produced by linear transductions from those components which arise out of non-linear transductions. The linear, or first order, responses come out at the input frequencies in the sinusoidal sum. The responses of non-linear elements are present as harmonic frequencies of the input frequencies, or as intermodulation frequencies which are additive combinations of two or more of the input frequencies (Victor, Shapley & Knight, 1977).

The major question asked in this paper is, how do the first order responses of retinal ganglion cells depend on contrast? If the retina were basically linear, or if non-linearities within the retina were connected in a relatively simple, serial manner, one would expect all the first order responses to be multiplied by the same constant factor as contrast increased (see Discussion). This is not what is found. Rather, the temporal transfer functions of retinal pathways are altered by contrast. This effect of contrast is seen in both X and Y cells of the cat retina, but the effect is larger in Y cells.

From our experimental observations, we have fashioned a model for the retina which includes the concept of a contrast gain control. The simplest adequate model is rather complicated, unfortunately. The mechanism seems to be equivalent to the shunting of a resistance or speeding up of a rate constant at higher contrast.

It may be thought that the mammalian retina is hopelessly complicated in detail, because each new series of experiments on the retina seems to unearth yet another complex non-linear mechanism. However, our work suggests that the situation may not be so bad, and that the contrast gain control may be directly related to the previously discovered non-linear subunits of cat Y cells (Hochstein & Shapley, 1976b; Victor et al. 1977). Our results indicate that the non-linear subunits and the contrast gain control have similar dependences on temporal frequency, spatial frequency, and spatial phase. So we have some hope that we can ultimately produce a single explanation for most of the complex, non-linear behaviour of cat retinal ganglion cells.
RETINAL CONTRAST GAIN CONTROL

METHODS

Recordings were made from optic tract fibres of anaesthetized (urethane) or decerebrate adult cats. The cat's e.g., e.g., blood pressure, core temperature, end-expiratory CO₂ and optics were monitored and maintained in the physiological range. Action potentials, recorded extracellularly with tungsten-in-glass microelectrodes, triggered a discriminator circuit which sent shaped pulses to a PDP 11/20 computer, which recorded their arrival time to within 0-1 msec.

Visual stimulation was accomplished with a cathode ray tube at a distance of 57 cm. The area of display was 20 cm × 20 cm, which spanned a visual angle of 20° × 20°. The mean luminance of the cathode ray tube was 10–20 cd/m². Spatial patterns were produced on the cathode ray tube with a specialized set of circuits (Shapley and Rossetto, 1976) to control the X-, Y-, and Z-inputs. The spatial patterns used in these experiments were standing sine gratings of arbitrary spatial phase and spatial frequency (oriented vertically) and rectangular spots of arbitrary dimensions and positions. The contrast of the pattern was modulated in time by a control signal from the PDP 11/20 computer. A control voltage of zero produced a uniform display at the mean luminance; when the control voltage passed through zero, the contrast reversed. The temporal modulation signal was either a single sinusoid, or a sum of nearly incommensurate sinusoids. In most of the experiments the signal was made up of eight sinusoids. When a single sinusoid formed the temporal modulation signal, neural responses were Fourier-analysed at the modulation frequency. When the sinusoidal-sum signal was used, the neural responses were Fourier-analysed at each of the input frequencies, as well as each of the second order frequencies (sums and differences of the input frequencies). The choice of the input frequencies allowed first and second order frequency kernels to be constructed from the Fourier coefficients (Victor et al. 1977; Victor & Knight, 1978). The input frequency sets used were chosen as described previously (Victor et al. 1977).

The receptive field of each optic tract fibre was mapped on a tangent screen. The receptive field centre was positioned in the center of the cathode ray tube display with a mirror, and the unit was classified as X or Y by a modified 'null test' (Hochstein & Shapley, 1976a; Victor et al. 1977). Then, the temporal modulation signal was placed under computer control to study dynamics of the response to many spatial patterns and contrasts. For each spatial pattern, several contrast levels were presented in interleaved runs. The contrast produced by each sinusoidal component was typically 0-0125, 0-025, 0-05, and 0-10 in successive runs. (Contrast = (Iₘₐₓ − Iₘᵢₙ)/(Iₘₐₓ + Iₘᵢₙ).) In other experiments, seven of the eight sinusoidal components each produced one contrast (typically 0-025), and the remaining sinusoid produced a higher contrast (typically 0-20). In this case, the frequency which was presented at higher contrast varied from run to run. In all cases, each contrast condition was presented several times, and the Fourier components from equivalent runs were averaged.

RESULTS

First, we will consider how we measured the first order frequency kernel of retinal ganglion cells. The spatial pattern used as a visual stimulus was modulated in time by a sum of sinusoids. The impulse train of the ganglion cell was Fourier-analysed. The amplitudes and phases of the responses at those temporal frequencies present in the stimulus make up the first order frequency kernel. Were the retinal transduction from light to nerve impulses linear, the first order frequency kernel would be the transfer function of that transduction. For a non-linear system, the significance of the first order frequency kernel is more complex, and the kernel's values may depend on the input signal used. However, for the particular input signal we used, the first order frequency kernel represents the transfer function of that linear system which best approximates the retinal pathway under study, where 'best' means best by the criterion of least squares. As the number of sine waves in the stimulus modulation signal becomes larger, the first order frequency kernel more and more closely
approximates the Fourier transform of the first order Wiener kernel (Victor & Knight, 1978).

First order kernels: the contrast effect

For a linear system, one would expect that the first order frequency kernel should grow proportionally with contrast. That is, the response amplitudes should double as contrast doubles, and the phases of the responses should remain the same. Any deviation from this behavior implies the presence of third order (or higher odd order) non-linearities (see Discussion). In fact, almost all optic tract fibres we studied showed evidence of this kind of non-linearity.

The first order frequency kernel was studied as a function of contrast, with a sine grating as the spatial stimulus, in thirty-one X cells (twenty-three on-centre, eight off-centre) and forty-one Y cells (thirty on-centre, eleven off-centre). In each case, the sinusoidal sum signal was presented at strengths separated by factors of two in interleaved episodes. The maximum contrast produced by each sinusoid was successively 0.0125, 0.025, 0.05, and 0.10. The input signal was composed of six or eight nearly incommensurate sinusoids so the root-mean-squared contrast in successive episodes was 0.025, 0.05, 0.10, and 0.20 for the eight-sinusoid experiments, and slightly less for the six-sinusoid experiments.

X cells. Data from a representative X cell are shown in Fig. 1. The unit was an on-centre X cell stimulated with a 0.2 cycles per degree (c/deg.) grating positioned to produce a peak first order response. At input temporal frequencies of 2 Hz and below, the response grew much less than linearly with input contrast. However, at the higher temporal frequencies of 8 Hz and 15 Hz, the response grew nearly proportionally to input contrast. Over nearly the entire frequency range tested, the phase shift of the responses advanced by about 0.2π radians as contrast increased over the range 0.0125 up to 0.1 per sinusoid.

To quantify the change in shape of the first order frequency kernel with contrast, we extracted two parameters from the data. One of these numbers was a ratio: the amplitude at 15 Hz divided by the amplitude at 0.5 Hz. The growth of this number as contrast increased quantified the change in the shape of the amplitude curve with contrast. The second parameter we examined was the phase shift of the response at 8 Hz. This number was a reliable index of the effect of contrast on the speeding-up of first order responses. Both these parameters are plotted vs. contrast in Fig. 2. Note that the amplitude ratio is plotted on a logarithmic axis, while the phase is graphed on a linear axis. This Figure summarizes the data for four X cells. The graph in Fig. 2A summarizes the data presented in Fig. 1. Fig. 2B represents the contrast effect in an off-centre X cell stimulated with a 0.25 c/deg. grating. Fig. 2C shows the contrast effect in an on-centre X cell stimulated by a 0.7 c/deg. grating. This is an interesting example of a large contrast effect in a cell which produced negligible second order non-linear responses. The graph in Fig. 2D is for an X cell stimulated by a 1 c/deg. grating. There was no contrast effect in this cell; the first order frequency kernels at contrasts from 0.0125 up to 0.1 were parallel. The flat graphs of the amplitude ratio and phase in Fig. 2D reflect this finding. Such behavior is what one would expect from a system which is basically linear.

One can see from Fig. 2 that the amplitude ratio of the responses at 15 and 0.5 Hz, and the phase advance at 8 Hz, are equivalent indices of the effect of contrast on the
first order frequency kernel. There are theoretical reasons for expecting this equivalence (DeGroot & Mazur, 1969). Later on in this paper we quantify the contrast effect solely in terms of the phase shift at 8 Hz, for reasons of convenience and

Fig. 1. First order frequency kernels as a function of stimulus contrast for an on-centre X cell. The input contrasts were 0.0125 (○), 0.025 (□), 0.05 (△) and 0.10 (●) per sinusoid. The spatial stimulus was a 0.2 c/deg. grating positioned to elicit a maximal linear response. In this and all other data illustrated, the mean luminance was 20 cd/m². Unit 11/4.

brevity. However, from Fig. 2 (and Fig. 4, below) one may conclude that the phase shift at 8 Hz is highly correlated with the change in shape of the amplitude curve.

Y cells. Typical Y cells showed a large effect of contrast on the first order kernel. This implied a greater amount of third order and perhaps higher odd order interactions in Y cells than in X cells. (This finding is not a trivial consequence of frequency-doubling by Y cells, for that is a manifestation of even order nonlinearities only.) A representative example is illustrated in Fig. 3. The unit was an off-centre Y cell. The spatial stimulus was a 0.25 c/deg. grating, positioned to produce a peak linear
response. At low temporal frequencies the response increased much less than proportionally with contrast. The responses at high temporal frequencies increased more than proportionally with contrast. As contrast varied over an eightfold range, the response at 0.5 Hz grew by a factor of about three while the response at 15 Hz grew by a factor of more than twenty. For input frequencies 1 Hz and higher, the phases of the responses advanced with increasing contrast. This advance amounted to 0.35 π radians at an input frequency of 8 Hz.

![Graph](image)

Fig. 2. Amplitude ratios and phase shifts of the first order frequency kernel as a function of stimulus contrast for three on-centre X cells (A, C, D) and one off-centre X cell (B). The spatial stimuli were sine gratings positioned to elicit a maximal linear response. The spatial frequencies used were 0.2 c/d (A), 0.25 c/d (B), 0.7 c/d (C) and 1.0 c/deg. (D). The ratio of the amplitudes of the response at 15 Hz to the response of 0.5 Hz (○), and the phase shift of the response at 8 Hz (△), are plotted as functions of the peak contrast of each sinusoid in the input signal. Units 11/4 (A), 8/4 (B), 18/4 (C) and 20/2 (D).

The amplitude ratio of responses at 15 Hz and 0.5 Hz, and the phase shift at 8 Hz, are plotted for three representative Y cells in Fig. 4. As in Fig. 2, these curves are summaries of first order frequency kernels obtained at four contrasts from 0.0125 up to 0.1 per sinusoid. Fig. 4A demonstrates the contrast effect for an on-centre Y cell stimulated with a 0.025 c/deg. grating. Fig. 4B represents the data graphed in Fig. 3. Fig. 4C summarizes the data from another on-centre Y cell which was
stimulated by a finer pattern, a 1.0 c/deg. grating. Fig. 4C illustrates the point that a contrast effect was observable in Y cells for any pattern which was effective in eliciting first order responses. The 1.0 c/deg. grating was near the spatial frequency resolution limit for this cell, yet at contrasts above 0.025 there was a clear contrast effect. The curves in Fig. 4 are steeper than those in Fig. 2, and this observation implies that the contrast effect was stronger in Y cells than in X cells.

Fig. 3. First order frequency kernels as a function of stimulus contrast for an off-centre Y cell. The input contrasts were 0.0125 (○), 0.025 (□), 0.05 (△) and 0.10 (●) per sinusoid. The spatial stimulus was a 0.25 c/deg. sine grating positioned to elicit a maximal linear response. Unit 12/3.

Several additional observations led us to the conclusion that the contrast effect was stronger in Y cells than in X cells. Besides the larger phase advances at high temporal frequencies, Y cells often produced more-than-proportional increase of response with contrast. Also the responses of Y cells to low frequencies of modulation sometimes even decreased in amplitude as contrast increased. Such very strong effects of an odd order non-linearity were rarely seen in X cells.
A graphical summary of the magnitude of the contrast effect is given in Fig. 5. Spatial sine gratings were most often employed as visual patterns in these experiments. Thus, in Fig. 5 the phase advance due to an eightfold increase in contrast of a spatial sine grating (from a contrast of 0.0125 up to 0.10) is plotted as the ordinate; the spatial frequency of the grating is the abscissa. These are the pooled data from the sixty-three retinal ganglion cells tested with eight sinusoids. Units which were tested with several gratings at different spatial frequencies are represented by several points, one for each spatial frequency investigated. It is clear that there is not a strong dependence of the contrast effect on spatial frequency but rather a gentle decline in the effect as spatial frequency increases. This decline in the contrast effect may be in part due to ineffectiveness of the fine gratings, those with spatial frequencies above 1.0 c/deg., to stimulate any of the retinal pathways strongly. Another fact illustrated by Fig. 5 is that on and off centre cells are affected similarly by the contrast of the stimulus.

The use of sine gratings as spatial stimuli was not crucial in our experiments on the contrast mechanism and its effect on the first order kernel. The same phenomenon was observed when the spatial stimulus was a spot or a bar (see Fig. 9). However, the grating stimuli did allow us to compare the spatial characteristics of the contrast mechanism with the previously studied non-linear excitatory subunits of Y cells (Hochstein & Shapley, 1976b; Victor et al. 1977). For cells which we stimulated with gratings over a wide range of spatial frequency, we could compare the strength of the second order responses with that of the contrast effect, as functions of spatial frequency. A graph of a representative experiment in a Y cell is displayed in Fig. 6. The measure of the contrast effect was the phase advance at 8 Hz over the contrast range from 0.0125 to 0.1. The strength of the second order responses was calculated by taking the root-mean-square of the second harmonic amplitudes.
measured at an intermediate contrast of 0.025. Fig. 6 demonstrates that both these measures have a similar dependence on spatial frequency, with a peak sensitivity near 0.25 c/deg.

![Graph](image)

Fig. 5. Strength of contrast gain control in sixty-three units tested under uniform conditions. The ordinate is a parameter to measure the strength of the contrast effect: the amount of phase advance of the 8 Hz response over the contrast range 0.0125-0.10 per sinusoid. The abscissa indicates the spatial frequency of the spatial sine gratings used, in cycles per degree. Data obtained with diffuse light are plotted at 0.01 c/deg. Filled symbols indicate X cells; open symbols indicate Y cells. Circles indicate on-centre units; triangles indicate off-centre units.

It is apparent from Fig. 5 that Y cells tended to show a greater phase advance than X cells. This is consistent with the fact that the amplitudes of the first order kernels showed a greater variation with contrast than did the kernels of X cells, as is evident in the examples of Figs. 2 and 4. A further illustration of the difference between X and Y cells is given in Fig. 7. This shows some of the data in Fig. 5, namely the points obtained with either a 0.2 or 0.25 c/deg. grating as a spatial stimulus over an eightfold range of contrast. A sine grating of 0.25 c/deg. was used as a stimulus because this spatial frequency is effective in stimulating both first and second order responses in Y cells. Fig. 7 shows the distribution of units with different degrees of phase advance. It is apparent that Y cells show a larger contrast effect than X cells on the average, though there is considerable overlap. The average phase advance for the nine X cells was 0.14 π radians (∼ 25°). The average phase shift for eighteen Y cells was 0.28 π radians (∼ 50°).
Responses to single sinusoids compared to responses to sums of sinusoids

The first order frequency kernels obtained at a series of contrasts were compared with transfer functions measured with the component sinusoids presented separately. One would expect these two measurements to be similar if the direct pathway from photoreceptors to ganglion cells had little odd order non-linearity. Conversely, odd order non-linearities should result in a dependence of the response at each particular frequency on the presence or absence of other sinusoidal components in the input signal.

![Graph showing phase shift and second harmonic strength as functions of spatial frequency](image)

**Fig. 6.** A comparison of the strengths of the contrast effect (●) and the second order response (▲) in an on-centre Y cell as a function of spatial frequency. The strength of the contrast effect is measured by the amount of phase advance of the 8 Hz response over the contrast range 0·0125 per sinusoid to 0·10 per sinusoid. The strength of the second order response is measured by the root-mean-squared amplitude of the responses at the second harmonics of the eight input frequencies. Unit 23/1.

In Fig. 8, responses of an on-centre X cell to single sinusoids and to a sum of sinusoids are compared. The temporal modulation signal was the sum of six sinusoids of equal amplitude and the following frequencies: 0·641, 1·10, 2·47, 5·37, 12·2, 21·4 Hz. The spatial pattern was a 1·0 c/deg. grating positioned to produce a maximal first order response. The amplitudes of the first order frequency kernels measured at different contrast levels (Fig. 8A) are nearly parallel. The phases of the response advance by at most 0·15 π radians as contrast increases. Therefore, in this X cell as in several other X cells, there was not very much effect of contrast on the first order frequency kernel. The responses elicited by each of the component sinusoids presented separately are shown in Fig. 8B. In general, the amplitudes and
phases of the responses measured in this way are similar to those measured when the sinusoids were presented simultaneously.

The situation is different for cells which show a strong contrast effect. A similar comparison between the first order kernel (measured with a sum of sinusoids) and the linear responses to single sinusoids is presented in Fig. 9. In this case the unit was a typical on-centre Y cell. The spatial stimulus was a bar, 0.5° wide, that covered the receptive field centre. The temporal modulation consisted of the six-frequency set used in Fig. 8A. The frequency kernels indicate the presence of a substantial contrast effect. At low temporal frequencies, response barely increased as contrast rose by a factor of eight. At high temporal frequencies, response increased nearly proportionally with contrast. There was a concomitant phase shift of 0.3 π radians at 5 Hz over the contrast range tested. The response amplitudes and phase shifts obtained by analysis of the responses to single sinusoids are graphed in Fig. 9B. These frequency responses differ in a crucial way from the first order frequency kernels. The amplitude functions showed much less of a change in shape with increasing contrast. Similarly, the phase advances obtained with single sinusoids were smaller than the phase advances obtained with a sum of six sinusoids of equal contrast.
The data of Fig. 9 suggest that the failure of response at some particular frequency to increase proportionally to contrast is not merely dependent on the strength of the input sinusoid at that frequency. This effect of contrast is also related to whether or not other sinusoids are present in the input stimulus. Therefore, the response to sinusoidal stimulation at a given temporal frequency depends not only on the strength of the Fourier component of the stimulus at that frequency, but also on the over-all power in the stimulus.

![Graph showing amplitude and phase for sum and single sinusoids.](image)

Fig. 8. A comparison of the first order frequency kernels obtained with a superposition of six sinusoids (A) and the responses to single sinusoids (B) in an on-centre X cell. The spatial stimulus was 1.0 c/deg. grating positioned to produce a maximal linear response. The contrasts used were 0.0125 (○), 0.025 (□), 0.05 (△) and 0.10 (●) per sinusoid. Unit 4/2.

The perturbation experiment

The results described above suggest that some function of the over-all power in the stimulus modifies the linear transfer properties of retinal ganglion cells. This effect is greatest in Y cells. One would expect that power at different temporal frequencies should have different degrees of influence on the linear transfer properties. The temporal frequency dependence of the contrast effect was explored as follows. A sinusoidal-sum signal was constructed with eight components (0.229, 0.473, 0.961,
but the amplitudes of the sinusoidal components were not all equal. Seven of the eight sinusoids produced a maximal contrast of 0.025, and the eighth sinusoid produced a contrast of 0.20. This procedure is illustrated in Fig. 10. The idea was that the low-contrast sinusoids could serve as probes of the transfer properties of the ganglion cell in question, by producing small linear responses over a wide frequency range. The eighth sinusoid contained nearly all of the

![Diagram](image)

**Fig. 9.** A comparison of the first order frequency kernels obtained with a superposition of six sinusoids (A) and the responses to single sinusoids (B) in an on-centre Y cell. The spatial stimulus was a 0.5 deg bar positioned to produce a maximal linear response. The contrasts used were 0.0125 (○), 0.025 (□), 0.05 (△) and 0.10 (●) per sinusoid. Unit 5/5.

power in the input signal. Its presence perturbed the responses at the other fundamental frequencies. The frequency of the high-contrast sinusoid was varied from episode to episode, so that the effect of power at a wide range of temporal frequencies could be observed. The perturbation of the linear responses to the low-contrast sinusoids was assayed by observing changes in phase shifts.

Results from two such experiments on Y cells are shown in Fig. 11. The unit of Fig. 11A was on-centre; the unit of Fig. 11B was off-centre. In each case, the spatial stimulus was a 0.12 c/deg. grating. In the left side of each panel, we have plotted the
phase shifts obtained in a perturbation experiment. The abscissa indicates the frequency of the perturbing sinusoid presented at a contrast of 0.20. The remaining seven assay sinusoids each produced a maximal contrast of 0.025. The ordinate measures the phase shift of the response at a given assay frequency. This phase shift is measured relative to the phase of the response obtained when all eight sinusoids were presented with a contrast of 0.025 per sinusoid. Three temporal frequencies (3.9 Hz, 7.8 Hz, and 15.6 Hz) were used as assay frequencies; the responses at these frequencies showed the largest shifts with contrast. Each curve contains only seven data points, because a frequency could be used as an assay frequency only when it was not the perturbing frequency. The right half of each panel shows the relative phase shifts obtained when all sinusoids were presented at equal strength. The abscissa is the contrast per sinusoid; the ordinate is the phase shift relative to the phase measured when all the sinusoids had an equal contrast of 0.025 per sinusoid.

It is apparent that the first order responses of both units showed a significant phase advance with contrast. The assay frequency at which the phase advance was maximal was 7.8 Hz for the unit in Fig. 11A, and 15.6 Hz for the unit in Fig. 11B. From graphs like those shown in Fig. 11, we determined which temporal frequencies were most effective in producing phase advances. There were the temporal frequencies in the range 4–15 Hz, the same temporal frequencies that were most influenced by contrast.

In Fig. 11B, the curves of phase advance as a function of perturbing frequency have roughly the same shape, independent of the temporal frequency used to assay the phase shift. In this regard, the data of Fig. 11A are somewhat different. The apparent effectiveness of various input frequencies in producing phase shifts
depends slightly on the assay frequency. The implications of this phenomenon for modelling the mechanism of the parametric dependence on contrast will be discussed below.

Fig. 11. Perturbation experiments in two Y cells. The first order frequency kernel was measured with a signal composed of seven sinusoids producing a contrast of 0.025 each, and one sinusoid producing a contrast of 0.20. In the left panels are shown the phase shifts at several assay frequencies (3.9 Hz (.), 7.8 Hz (■), and 15.6 Hz (▲)) as a function of the perturbing frequency. In the right panels, this comparison was made for input signals consisting of equal-strength sinusoids, producing contrasts of 0.025, 0.05, and 0.10 per sinusoid. In all cases, the spatial stimulus was a 0.12 c/deg. grating positioned to produce a peak linear response. Units 7/7 (A) and 7/9 (B).

Second order kernels and the perturbation experiment

The results of the perturbation experiments indicate the dynamic characteristics of the contrast mechanism. We suspected that there was a relation between the non-linear subunits, which generate the second order responses of Y cells, and the contrast mechanism. Therefore, we compared the temporal frequency dependence of the contrast effect with that of the second order frequency kernels (Victor et al. 1977).
This was done for the Y cells of Fig. 11. Contour maps of the amplitudes of the second order kernels are shown in Fig. 12A, B.

The conventions for the display of the second order frequency kernel as a contour map are as follows: The co-ordinates in the plane range independently over the frequencies in the input stimulus. The height of the surface at a point at \((F_1, F_2)\) indicates the amplitude of the response at the frequency \(F_1 + F_2\). This is the amplitude of the second order frequency kernel, \(K_2(F_1, F_2)\). Similarly, the height of the surface at a point \((-F_1, F_2)\) indicates the amplitude of the response at \(F_2 - F_1\), which is \(K_2(-F_1, F_2)\). For each pair of temporal frequencies \(f_1\) and \(f_2\) in the input signal, these values were determined at points \((\pm f_1, f_2)\) by Fourier analysis of the impulse train. For combinatorial reasons, the Fourier component at \(2f_1\) was doubled to obtain \(K_2(f_1, f_2)\). Also, the values \(K_2(-f_1, f_2)\), which correspond to an output frequency of zero, are not defined. These values, along with the values \(K_2(\pm f_1, F_2)\) for frequencies \(F_2\) and/or \(F_1\) not in the stimulus were interpolated by a standard two-dimensional cubic spline. The resulting surface was plotted as a contour map, with a vertical scale in which one contour line indicates one spike per second. The tickmarks point downhill. The identity \(K_2(F_1, F_2) = K_2(F_2, F_1)\) results in a line of mirror symmetry running on a 45° angle through the sum frequency region (the upper half of the map). This is the line \(F_1 = F_2\), the line of second harmonics. Similarly, the relationship \(K_2(-F_1, F_2) = K_2(F_2, -F_1)\) creates a line of mirror symmetry running at a 45° angle through the difference frequency region (the lower half of the map). This line, \(F_2 = -F_1\), is the line of zero output frequency. It is perpendicular to the first symmetry line, the line of second harmonics.

The amplitudes of the second order kernels for these Y cells are typical in having a peak at intermediate input frequencies (4–8 Hz) and a steep roll-off at low and high input frequencies. The best frequencies for producing second order responses are also those frequencies which produce most contrast effect, as measured by the phase advances in Fig. 11. This is illustrated for the second harmonic frequencies from the second order frequency kernels in Fig. 12. The graphs of the second harmonic amplitudes \(v\), input frequency are located above the appropriate contour map.

**Parametric dependence on contrast: independence of spatial phase**

It is important to know whether the shape of the first order frequency kernel depends on retinal contrast or rather on the size of the first order responses themselves.

We could control independently retinal contrast and size of response by variation of the spatial phase of the grating used as a visual stimulus. By re-positioning a grating of a given contrast at several spatial phases, we varied first order responses without changing retinal contrast. If the shape of the first order frequency kernel (aside from absolute magnitude) depended on the spatial phase, we could conclude that the size of the responses themselves, rather than contrast alone, is the cause of this change of shape. This was not the case (as shown below), so we concluded that response size per se is irrelevant to the contrast effect. Alternatively, we could vary retinal contrast but maintain an approximately constant size of response, by comparing the frequency kernel of a low contrast grating at a peak spatial phase with that of a high contrast grating near the spatial phase for a null response. This is another test of whether response size rather than retinal contrast determines the shape of the frequency kernels, for under that hypothesis two such kernels should be similar in shape. In fact, these frequency kernels were different in shape. Thus we concluded that it is retinal contrast which alters the shape of the first order frequency kernels.

Results from two such experiments are shown in Fig. 13. The data of Fig. 13 were
obtained from an on-centre X cell by varying the spatial phase of a 0.1 c/deg. grating. The size of the response to the grating presented with a temporal modulation of 0.10 per sinusoid was reduced by a factor of approximately ten by moving the grating from a peak position to a spatial phase near the null position. But

Fig. 12. Second order frequency kernels and slices along the second harmonic diagonal of the two Y cells in Fig. 11. The frequency kernels were obtained with a temporal stimulus consisting of eight sinusoids, each with a contrast of 0.025, and a spatial grating of frequency 0.12 c/deg. positioned to produce a maximal linear response. (This configuration was the reference condition for the perturbation experiments.)
Despite the change in the absolute magnitude of the response, the shape of the first order frequency kernel was essentially unaltered. The comparison of the high contrast grating near its null with a low contrast grating (0.0125 contrast) at its peak position is also instructive. The responses to these two stimuli were roughly equal on the average. Nevertheless, there is a striking difference between the shapes of the two frequency response curves. Previously we have used the phase shift at 8 Hz as an assay of the contrast effect. In this experiment too this measurement is instructive. The phase shifts for the two responses to the grating at 0.10 contrast were $-0.27\pi$ radians (peak) and $-0.35\pi$ (near null). The phase shift for the low contrast grating was $-0.53\pi$.

![Graph](image)

Fig. 13. A comparison of the effects of changing the contrast and of changing the spatial phase of a sine grating on the first order frequency kernel. For the on-centre X cell of A, data were obtained with a 0.10 c/deg grating positioned to produce a maximal linear response with contrasts of 0.0125 (○) and 0.10 (△) per sinusoid. When the grating was positioned close to the null position, a contrast of 0.10 per sinusoid (△) was used. For the on-centre Y cell of B, the spatial stimulus was a 0.25 c/deg grating. In a position of peak linear response, contrasts of 0.0125 (○) and 0.05 (△) per sinusoid were used. Near the null position, the temporal signal produced a contrast of 0.05 per sinusoid (△). Units 24/2 (A) and 8/6 (B).

Essentially the same behaviour is shown by the data obtained from an on-centre Y cell and presented graphically in Fig. 13B. Here, the spatial phase of a 0.25 c/deg grating at a contrast of 0.05 was varied to attenuate the first order response by a factor of four. Only the magnitude of the first order frequency kernel, not its shape, varied with spatial phase. However, attenuation of the contrast of the grating by a factor of four, to 0.0125, had a profound effect on the shape of the first order frequency kernel. The temporal frequency at which the first order response was maximal shifted from 8 Hz down to approximately 2 Hz. In this unit again, phase shift at 8 Hz was a reliable assay of the contrast effect. The phase shifts for the high contrast gratings were $-0.59\pi$ (peak) and $-0.60\pi$ (near null). For the low contrast grating the phase shift was $-0.91\pi$. 
DISCUSSION

The results reported here lead to the hypothesis of a distinct non-linear mechanism in the cat retina. This mechanism adjusts the sensitivity and dynamic characteristics of the retina contingent on the average contrast of visual stimuli presented to the retina. This mechanism affects the first order responses of both X and Y cells, though it has a stronger effect on Y cells. We will refer to this non-linear mechanism as the contrast gain control or the contrast mechanism. Crucial features of the contrast gain control are as follows: (1) it affects phase shifts at high temporal frequencies as well as amplitudes at low temporal frequencies, (2) it allows energy at one temporal frequency to affect amplitude and phase shift at other frequencies, (3) it is relatively insensitive to slow modulation frequencies and the mean light level and (4) it is independent of spatial phase and not greatly dependent on the spatial frequency of the visual pattern.

The contrast effect is not trivial

First we show that the contrast effect must reflect internal properties of the retina. What must be excluded is the hypothesis that the change in shape of the first order frequency kernels is the consequence of a static saturation. This hypothesis is excluded because: (1) as contrast increases, the responses which were already large get even larger, (2) a major effect of contrast is on phase shift of the first order responses and (3) retinal contrast, rather than response size, controls the shape of the first order responses. The last fact also excludes the possibility that a peculiarity of the spike-generating mechanism (or indeed any other transduction after final spatial pooling) is responsible for the effects we have reported.

We can also exclude light or dark adaptation as an explanation of the effects of contrast. Fast temporal frequencies (4–15 Hz) were most effective in exerting an influence on the transfer characteristics (see Fig. 11). Thus any mechanism which is primarily sensitive to mean level, such as the gain control of light adaptation (Enroth-Cugell & Shapley, 1973), is ruled out.

Models for the effect of contrast on the first order frequency kernel

On general grounds, a change in shape of the first order frequency kernel with contrast requires the existence of third order or higher odd order interactions.

One can calculate directly the frequency kernels for a system whose non-linearities are of finite order. This class of systems does not include systems that contain sharp non-linearities. Nevertheless, the calculation is useful because it demonstrates the minimal formal requirements for a contrast effect on the first-order frequency kernel. Suppose for example that the non-linearities in a system are of order no higher than three. In this instance, the first order frequency kernel has the form

\[ K_1(f_j) = a_1 g_1(f_j) + \frac{1}{\pi} \sum_{k=1}^{\infty} a_k g_k(f_j, f_k, -f_j) + \frac{1}{\pi} \sum_{k=1}^{\infty} b_k g_k(f_j, f_k, -f_j). \]  

(1)

In this equation, \( a_1 \) and \( f_j \) are the amplitude and frequency of the \( j \)th sinusoidal component in the input signal. The functions \( g_k \) depend only on the system under study, and are Fourier transforms of the Volterra kernels. In particular, the function \( g_1 \) is the small signal transfer function.
The first order frequency kernel (eqn. (1)) depends not only on \( g_1 \), but also on \( g_2 \). (In fact, all higher order odd order interactions \( g_2, g_4, g_6, \ldots \) would influence the first order frequency kernel, were they present in the system.) Since the coefficients of \( g_2 \) are cubic in the input amplitude, the contributions of these terms will not scale linearly with input amplitude. The result is that the first order frequency kernel \( K_1 \) may change shape as input amplitude increases, for it is the sum of components that scale as different powers (first and third powers) of the input amplitude.

Yet there are many non-linear systems which contain high odd order non-linear interactions and which nevertheless do not show such a contrast effect. These non-linear systems are those that consist of an arbitrary linear filter, \( L_1 \), followed by a static non-linearity \( N \), followed by a second linear filter, \( L_2 \) (Fig. 14A). The first order frequency kernel for such a system has the approximate form

\[
K_1(f_1) = c_1 \hat{L}_1(f_1) \hat{L}_2(f_1),
\]

where \( \hat{L}_1 \) and \( \hat{L}_2 \) are the transfer functions of the two linear filters. This formula follows from the close relationship (Victor & Knight, 1978) of the frequency kernels and the Fourier transform of the Wiener kernels (Wiener, 1958). In eqn. (2), \( c_1 \) is a real number that depends on the input signal power and on the characteristics of the non-linear element \( N \). Eqn. (2) fails to explain the qualitative features of our ganglion cell data. This is because changing \( c_1 \) can only change the over-all size rather than the shape of the frequency kernel.

Fig. 14. Simple model non-linear systems rejected as candidates to explain the contrast gain-control. In A, the input signal is transformed in sequence by a linear filter, \( L_1 \), a static nonlinearity, \( N \), and a second linear filter, \( L_2 \). The model of B is a parallel combination of two ‘sandwiches,’ but with different linear and non-linear components.

We are therefore forced to consider non-linear systems more general than a linear/non-linear/linear sandwich, for example a network that consists of a parallel combination of such sandwiches. A hypothetical system consisting of two linear/non-linear/linear sandwiches in parallel combination (Fig. 14B) can produce a contrast effect in some ways similar to what we have observed, provided that the prefilters \( (L_2, L'_2) \) and the static non-linearities \( (N, N') \) have different characteristics.
A particularly attractive 'multiple sandwich' hypothesis is that one path corresponds to the classical 'centre' mechanism and the other corresponds to the classical 'surround'. The predictions of this model are at variance with major qualitative features of our data: (1) the dependence of the shape of the first order frequency kernel on input contrast persists for spatial sine gratings of high spatial frequency, which do not stimulate the classical surround substantially (Figs. 2C, 4C, and 5) and (2) since the contrast effect is independent of spatial phase (Fig. 13), it is therefore independent of the degree of net stimulation of either the center or the surround of the unit in question.

Fig. 15. A two-input non-linear system contemplated as a model for the contrast gain control. The filter L transforms its direct input, $I_D$, linearly, provided that the contrast input, $I_C$, is fixed. $I_C$ is extracted from the input signal by the nonlinear network, C, whose output contains only even order components. The signal $I_C$ alters the characteristics of L in a dynamic fashion.

Two-input models. The spatial phase invariance of the shape of the first order frequency kernel is a highly constraining fact. Therefore we developed a model that contains this feature in its initial formulation. This model is shown in Fig. 15.

We consider the classical centre/surround mechanisms of a ganglion cell's receptive field to constitute the approximately linear filter, L, and its direct input, $I_D$. This direct pathway is approximately linear in space as well as time, so its response to inputs along this pathway varies sinusoidally with the spatial phase of the stimulus. We postulate that its deviations from linearity are gentle, inessential ones that can be ignored in the stimulus range used in these studies. However, the linear element receives another input, $I_C$. The signal $I_C$ is produced by the contrast gain control C. We hypothesize that C measures the contrast of the stimulus at many separated points in the visual field over a region at least as large as the conventional centre and surround. In this way, the signal $I_C$ can be independent of the spatial phase of a grating stimulus. (The spatial phase independence would not apply to spatial gratings so low in spatial frequency that they are equivalent to diffuse light.) Since contrast, rather than intensity, is measured by $I_C$, the network C which produces $I_C$ must contain an even order non-linearity. We propose that, given a fixed contrast signal $I_C$, the filter in the direct pathway transforms its direct input, $I_D$, in a linear
manner. This formalizes the concept that \( L \) is a basically linear filter, but that the filter characteristics are parametric in local retinal contrast.

The interaction of \( I_C \) and \( I_D \). The contrast signal \( I_C \) may be thought of as primarily a steady level which may have a small modulated component. If the contrast gain control signal \( I_C \) had a large modulation, the phase shift curves in Fig. 11 would depend strongly on the assay frequency used, and they do not.

The effect of contrast on both amplitude and phase rules out hypotheses that the contrast signal \( I_C \) affects only the gain or only the phase shift of \( L \). We need to hypothesize a mode of action of \( I_C \) on \( L \) that alters both amplitude and phase of the first order frequency kernel. As contrast increases, the responses at the high frequencies are enhanced relative to the responses at low frequencies. This suggests that the effect of contrast may be to decrease the number of effective low pass stages in the linear filter of the direct pathway, \( L \). This would produce a phase advance at frequencies above the corner frequency of the low pass filters involved. In addition, elevated contrast levels signalled by \( I_C \) might act to shorten time constants in the direct pathway. This would produce a phase advance over those frequencies between the corner frequency of the high frequency roll-off at low contrast and the corner frequency at high contrast.

However, these mechanisms do not explain why the amplitudes of the response at low temporal frequencies (2 Hz or below) grow much less than proportionally with contrast. An increase in the effective number of high pass stages in the direct pathway filter, \( L \), might explain this behavior. Furthermore, such a parametric dependence on the contrast signal, \( I_C \), will also cause a phase change at low temporal frequencies, in agreement with the data. Another possible model that could create the necessary high and low frequency changes with a single mechanism is one in which the contrast signal, \( I_C \), increases the amount of feedback in a ‘self inhibiting’ loop within the filter \( L \).

Characteristics of the contrast network

The fact that the contrast gain control, \( C \), measures contrast, rather than illumination, implies that its response must be even order in light intensity. A contrast-reversed pattern has to produce in \( C \) the same response as the original pattern. Thus, \( C \) must produce only even order non-linear responses. But the fact that relatively fine gratings are adequate stimuli for this network implies that the spatial pooling which occurs before the even order non-linearity must be limited in extent. In addition, the fact that the output of this network is insensitive to spatial phase implies that its response must contain contributions from many spatial pools. (This is precisely the argument used to deduce the existence of subunits in the receptive field of \( Y \) cells from the existence of a spatial phase-invariant frequency-doubled response (Hochstein & Shapley, 1976b; Victor et al. 1977).

It is likely that the non-linear subunits involved in generation of the contrast signal are the same subunits that lead to the characteristic frequency-doubled responses of \( Y \) cells (Hochstein & Shapley, 1976b; Victor et al. 1977). The perturbation experiments show that the temporal properties of the input to the contrast mechanism are similar to the temporal properties of the second order frequency kernels of \( Y \) cells. Also, the spatial characteristics of patterns which are effective
stimuli for the contrast mechanism and the Y cell non-linear pathway are similar, as was shown directly in Fig. 6. We cannot exclude now the possibility that the subunits involved in C are distinct from those involved in generating the non-linear excitatory response of Y cells. However, it must be true that these two non-linear mechanisms have similar spatial and temporal characteristics.

Therefore, a plausible though speculative hypothesis is that there is one major essential non-linearity in the retina. It drives the Y cells directly and excites them. By another pathway, it suppresses the responses of all ganglion cells to slow changes but speeds up and perhaps boosts the responses to high temporal frequencies of modulation.

Relating to previous work

The contrast mechanism required by our findings can be invoked to provide an explanation for many other seemingly diverse results. For example, the responses of cat ganglion cells to square wave illumination of an ascending series of intensities become more transient and have a reduced latency (Cleland & Enroth-Cugell, 1970). These effects could be accounted for by the automatic gain control responsible for light adaptation (Cleland & Enroth-Cugell, 1970; Enroth-Cugell & Shapley, 1973). However, the latency changes and changes in dynamics suggest that the contrast mechanism may be even more important in the production of these effects. In comparable experiments performed in our laboratory, we have observed a similar phenomenon in the responses of X and Y cells to square-wave contrast reversal of sine gratings. As contrast was increased, the square-wave responses became more transient and the latency decreased (Hochstein, S., Kaplan, E., Shapley, R. M. and Victor, J. D., unpublished results). The latency decreased about twice as much for Y cells as for X cells. The magnitude and characteristics of this effect suggest that the contrast mechanism is probably involved. The fact that the contrast gain control affects Y cells more than X cells may explain why the square-wave responses of Y cells tend to be more transient than those of X cells (Cleland, Dubin & Levick, 1971) when high contrast stimuli are used. The square wave responses of X and Y cells can be quite similar at lower contrast (Hochstein & Shapley, 1976a).

The idea of two gain controls originated with Werblin's work on adaptation in the mudpuppy retina (Werblin, 1972; Werblin & Copenhagen, 1974). The flux gain control affected bipolar cells and therefore every other element in the retina proximal to the bipolar. The gain control for change, what we refer to here as the contrast gain control, was born in the inner plexiform layer and affected amacrine and ganglion cell responses. The mudpuppy experiments were done with a spinning windmill pattern. This windmill was the spatio-temporal visual stimulus used to excite the contrast gain control. Other workers have used similar stimuli on cat ganglion cells. Using a spinning windmill or a rotating radial grating, they have found evidence for a 'suppressive surround' in X, Y, and W ganglion cells (Jakiela, 1978, and personal communication; Cleland & Levick, 1974). This 'suppressive surround' is, we believe, another name for the contrast gain control. In another paper it will be shown that the characteristics of the 'suppressive surround' are the same as those of the contrast mechanism (Shapley, R. M. and Victor, J. D., in preparation).
We thank F. Ratliff for his advice and encouragement. This work was supported by grants from the U.S. National Eye Institute nos. EY 188, EY 1428 and EY 1472. One of us (R.M.S.) was supported by a Career Development Award from the National Eye Institute. Also, we thank Dr William Scott of Hoffmann-LaRoche for his gift of diallyl-bis-nortoxiferine used in the experiments.

REFERENCES


