J. Physiol. (1981), **318**, *pp.* 161–179
With 6 text-figures
Printed in Great Britain

HOW THE CONTRAST GAIN CONTROL MODIFIES THE FREQUENCY RESPONSES OF CAT RETINAL GANGLION CELLS

BY R. M. SHAPLEY AND J. D. VICTOR
From the Rockefeller University, New York N.Y. 10021 U.S.A.

(*Received 15 August 1980*)

SUMMARY

1. A model is proposed for the effect of contrast on the first-order frequency responses of cat retinal ganglion cells. The model consists of several cascaded low pass filters ('leaky integrators') followed by a single stage of negative feed-back.

2. Values of time constants and gain of the components in this model were chosen to approximate (with least-squared deviation) experimentally measured first-order frequency responses. In the experiments used for the analysis, the visual stimulus was a sine grating modulated by a sum of sinusoids.

3. For both X cells and Y cells, the over-all gain and the time constants of the cascade of low pass filters were insensitive to contrast.

4. In all cells, the gain-bandwidth product of the negative feed-back loop was markedly increased with increasing contrast.

5. The effect of stimulation in the periphery of the receptive fields on the first-order frequency response to a centrally placed spot was identical to the effect of increasing contrast in the grating experiments. In all cases, the gain-bandwidth product of the negative feed-back loop was the only model parameter affected by peripheral stimulation.

6. A similar effect of non-linear summation was investigated for two bars located in the receptive field periphery.

7. This analysis of the contrast gain control mechanism is compared with other models of retinal function.

INTRODUCTION

The research reported in this paper was undertaken in order to understand how contrast modifies the temporal tuning of cat retinal ganglion cells. We have previously demonstrated that there is an effect of contrast on the frequency responses of 'linear' and 'non-linear' pathways in the cat retina (Shapley & Victor, 1978, 1980). One effect of contrast is the relative enhancement of the amplitudes of responses to higher temporal frequencies at higher contrasts. This occurs because the amplitudes of response to higher temporal frequencies are often approximately proportional to stimulus contrast while the amplitudes of response to low temporal frequencies (< 2 Hz) usually increase less than proportionally to contrast. Accompanying the relative attenuation of low frequency amplitudes as contrast increases is a phase advance of the responses to frequencies of modulation in the 4–16 Hz band. From

experiments on the spatial frequency dependence, the spatial phase dependence, and the temporal frequency dependence of these effects of contrast, we have inferred that a specific retinal mechanism, the contrast gain control, is responsible for the effects (Shapley & Victor, 1978, 1979).

The question we now seek to answer is, precisely how does the contrast gain control adjust the dynamics of the retina? Another way of stating the problem is to ask whether the amplitude suppression (at low temporal frequency) and the phase advances (at mid-range frequencies) are due to a single action of the contrast gain control, or whether they are two separate manifestations of its effect on different parts of the retina. One of the main results of the paper is that the contrast gain control probably has only a single site of action.

METHODS

Recordings were made from optic tract fibres of anaesthetized (urethane) adult cats. The cat's e.c.g., e.e.g., blood pressure, core temperature, end-expiratory CO_2 and optics were monitored and maintained in the physiological range. Action potentials, recorded extracellularly with tungsten-in-glass micro-electrodes, 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, spanning a visual angle of 20 deg × 20 deg. Spatial patterns were produced on the cathode ray tube with a specialized set of circuits (Shapley & Rossetto, 1976) to control the X-, Y-, and Z-inputs. The spatial patterns used in these experiments were standing sine gratings of adjustable 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 a sum of eight nearly incommensurate sinusoids. The eight frequencies used in the experiments reported here were: 0·219, 0·458, 0·946, 1·923, 3·876, 7·782, 15·594, 31·219 Hz. The neural responses were Fourier-analysed at each of the input frequencies, to construct the first-order frequency kernel. The input frequency sets were chosen as described previously (Victor & Shapley, 1979); in all cases the procedure of varying initial phases to remove high-order overlaps was applied (Victor & Shapley, 1980).

Each experimental episode lasted 32·768 sec. With phase averaging, eight of these episodes were averaged together. Thus, each frequency response is the result of the analysis of over 4 min of neural activity. Typically, the standard error of the response amplitude was less than 2 impulses/sec. and the standard deviation of the phase was less than 0·05 π radians.

The receptive field of each optic tract fibre was mapped on a tangent screen. The receptive field centre was positioned in the centre 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, 1976). Then, the display was placed under computer control to study the response to a series of spatial patterns and contrasts.

The 'contrast experiment' protocol consisted of a grating presented at several contrasts 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 $= (I_{max}-I_{min})/(I_{max}+I_{min})$, where I_{max} is the maximum luminance and I_{min} the minimum luminance of the stimulus.)

The 'summation experiment' protocol consisted of four basic episodes. In the first episode type, one region was modulated by the sum of sinusoids and the second region was held at the mean luminance. Secondly, both regions were modulated together, in phase. Thirdly, the second region alone was modulated and the first region was held at the mean luminance. Fourthly, both regions were modulated in antiphase. The interleaving of episodes tended to compensate for any trends in the data due, for instance, to fluctuations in the sensitivity of the retina. As in the contrast experiment, each type of episode was repeated 8 times to increase the signal-to-noise ratio and to apply the algorithm for removal of high-order overlaps by phase averaging (Victor & Shapley, 1980).

The summation experiments were performed in two ways. To study centre/periphery interactions, one region consisted of a small spot positioned over the receptive field centre, and the second region consisted of four large blocks separated from the centre stimulus (Shapley & Victor, 1979). To study interactions in the periphery, the two regions were rectangular bars on opposite sides of the receptive field centre.

Curve fitting procedure

We fitted the measured first-order frequency response using the formula in eqn. (1) for several levels of stimulus contrast (denoted C). For each contrast, the best fitting parameters, A, τ_L, τ_H, and k in eqn. (1) were obtained by a nonlinear least-squares procedure (Fletcher & Powell, 1963). The function to be minimized, R was the weighted sum of the squares of the distances in the complex plane between the logarithm of the observed first-order frequency response, and the logarithm of the value given by eqn. (1). The individual squared distances between experimental and fitted values of the logarithm of the kernel were weighted proportionally to the observed amplitudes. This weighting expresses the greater reliability of the larger responses. Thus, the function we minimized was

$$R(A, \tau_L, \tau_H, k, N_L, N_H) = \sum_{j=1}^{8} w_j |\log K_1(f_j) - \log cG(f_j, A, \tau_L, \tau_H, k, N_L, N_H)|^2$$

where

$$w_j = \frac{|K_1(f_j)|}{\sum_{m=1}^{8} |K_1(f_m)|}$$

We will call $R(A, \tau_L, \tau_H, k, N_L, N_H)$ the residual.

Each set of data was fit separately by the following procedure. A data set consisted of first-order frequency responses measured at several contrasts, all other stimulus variables held fixed. The integers N_L and N_H were fixed for each data set. These parameters were chosen by first determining an N_L^{opt} and N_H^{opt} that provided the minimum residual individually for each contrast. This procedure always yielded $N_H^{opt} = 1$. Furthermore, the dependence of the residuals on N_L was gentle. Therefore, a uniform compromise value of N_L could be chosen that fit the frequency kernels at all contrasts well, and N_H was fixed at 1, as shown in Fig. 2. Then the non-linear least-squares procedure was applied to the entire data set. We checked this fitting procedure by comparing the residual R, calculated with the uniform compromise value of N_L, with an optimal residual R^{opt} which was calculated by letting N_L seek its optimal value to minimize the residual.

Results of typical calculations are shown in Table 1. The residuals, R, which are typically about 0·05, indicate that the fitted functional forms deviate in phase by no more than $0{\cdot}07\,\pi$ radians, or in amplitude by no more than 25% from any measured point. The small values of $R - R^{opt}$ indicate that forcing uniform values of N_L leads to only a small increase in the residuals. A comparison of the measured frequency kernels and the values fitted with a uniform N_L is illustrated in Fig. 1. It is evident that the functional form (1) provides a good phenomenological description of the amplitudes and phases of the responses, at both high and low levels of stimulus contrast.

RESULTS

The effect of contrast

First, we present typical results on the effect of contrast on the first-order frequency response of a retinal ganglion cell. Then we will introduce our model for the first-order pathway of the ganglion cell and will show how the responses of the model were fit to the cell's responses.

The effect of contrast was measured in a standard way. The visual stimulus was a spatial sine grating modulated in time by a sum of eight sinusoids. In a typical experiment like the one which yielded the results in Fig. 1, the contrasts used were 0·0125, 0·025, 0·05 and 0·1 per sinusoid. This means that the average contrasts (root-mean-square) were 0·025, 0·05, 0·1 and 0·2 for these four contrast levels. In Fig. 1 only the results for the extremes of the contrast range are shown. The results were

obtained from an on-centre Y cell. The spatial stimulus was a 0·25 cycle/deg. sine grating positioned to give a maximal first-order response.

The experimental results in Fig. 1 show clearly the typical effects of contrast on the first-order frequency response. The peak of the amplitude curve shifted to the right (higher temporal frequency) at higher contrast, and the high-frequency responses were much larger compared to low-frequency responses. The lower panel of the figure is evidence of a phase advance at higher contrast. Note that in Fig. 1

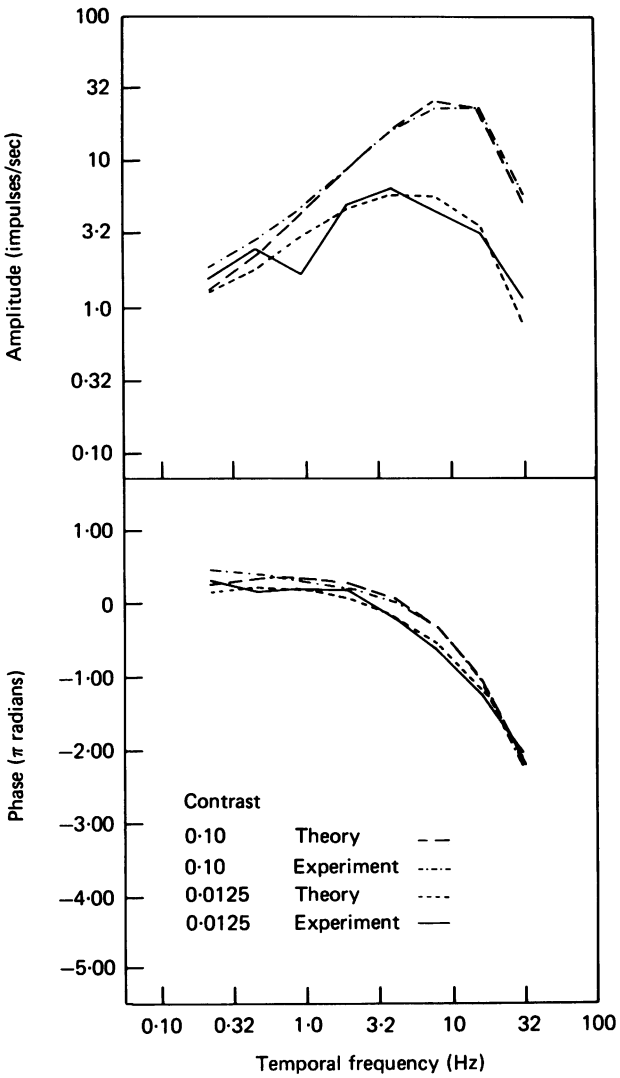

Fig. 1. A comparison of measured first-order responses to a grating at high and low contrasts, with the response of a lumped model. An on-centre Y cell was stimulated with a 0·25 cycles/deg grating, positioned to elicit a maximal first-order response. The eight-sinusoid sum produced a contrast per sinusoid of 0·0125 (——) and 0·1 (–·–·). The first-order frequency responses of a model transducer (equation 1) are shown for these contrast levels, 0·0125 (. . .) and 0·1 (– – –). The parameters of the model are given in Table 1. Mean luminance in this experiment and in all subsequent figures was 20 cd/m^2. Unit 8/3.

only two of the curves in each panel are plots of experimental data. The other two curves in each panel are derived from calculation with our proposed model and are labelled Theory. We now proceed to discuss the model which generated these theoretical functions.

A model for the first-order pathway

To analyse the dependence of retinal transduction on contrast, a functional form for the first-order frequency response was chosen, and the dependence of the form's parameters on contrast was investigated. The functional form corresponds to a lumped model with simple components (Fig. 2). The model transducer has an over-all

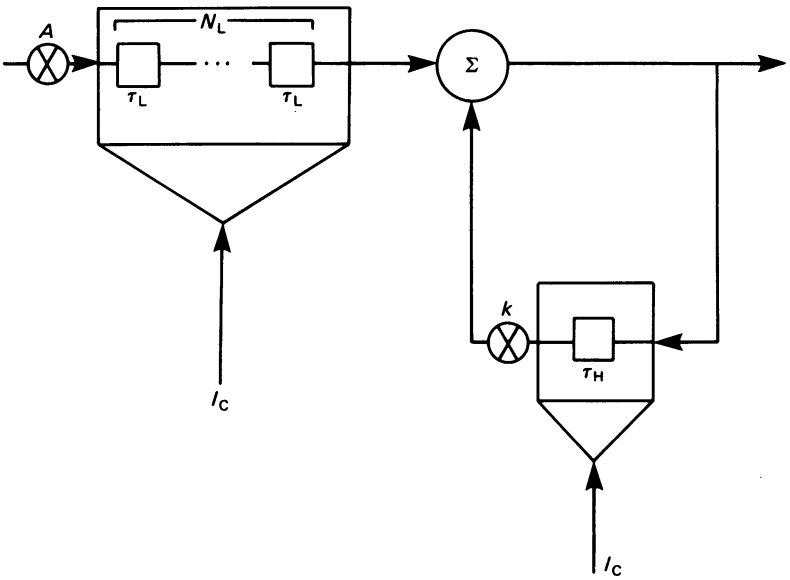

Fig. 2. A lumped model for the dependence of the first-order response of a ganglion cell on the contrast signal, I_C. The model consists of a gain stage of A, N_L low-pass stages, and N_H high-pass stages. Each low-pass stage has time-constant τ_L. Each high-pass stage (only one is illustrated) consists of a feed-back loop containing a low-pass stage of time constant τ_H connected with feed-back strength k. The parameters A, τ_L, τ_H, and k are permitted to vary with contrast.

gain of A, and stereotyped low-pass and high-pass components. The low-pass component consists of N_L stages, each of time constant τ_L. The high-pass component consists of N_H stages. Each of these stages consists of a low-pass filter of time constant τ_H in a feed-back loop with feed-back strength k. Thus, the transfer function of the model transduction is

$$G(f, A, \tau_L, \tau_H, k, N_L, N_H) = A \cdot \left(\frac{1}{1+2\pi i f \tau_L}\right)^{N_L} \cdot \left(\frac{1}{1+k/(1+2\pi i f \tau_H)}\right)^{N_H}. \quad (1)$$

The initial basis for using this kind of model was previous work on the transfer characteristics of photoreceptors and retinal interneurons (Baylor, Hodgkin & Lamb, 1974; Pasino & Marchiafava, 1976; Toyoda, 1974; Naka, Marmarelis & Chan, 1975). It is well known that the transfer functions of photoreceptors are well fit by a cascade

of low-pass elements of the type we propose in eqn. (1) (Baylor *et al.* 1974). The transfer characteristics of bipolar cells and ganglion cells reveal the presence of high pass stages in the proximal retina (Toyoda, 1974; Naka *et al.* 1975). Our idea that this high pass stage (or stages) might result from neural feed-back was an inference based on retinal anatomy (e.g. Dowling & Boycott, 1966).

Dependence of the model parameters on contrast. We next examine the pattern of the dependence of the parameters of the model transduction (1) on input contrast. Typical results are shown in Table 1.

Low pass stages. For each unit studied, the time constant of the low-pass stages, τ_L, is virtually independent of contrast. In the ganglion cells of Table 1, τ_L varies by about 10% as contrast increases, and shows no consistent pattern of dependence (it increases monotonically with contrast in unit 22/16, it decreases monotonically in unit 8/4, and shows a maximum in units 22/13 and 8/3). Across our population of units, the number of low-pass stages, N_L, and their time constants, τ_L, vary by a factor of about 2. However, their product $N_L \tau_L$ is remarkably constant. This product, which indicates the total delay of the low-pass stages, has a value of about 40 msec at all contrast levels, and in all units. Thus, the low-pass component of the transduction is not affected by the contrast gain control.

Gain. The overall gain, A, varies significantly from cell to cell. In different cells the dependence of A on contrast shows different patterns. For example, in unit 8/3 of Table 1, A is essentially independent of contrast; in units 22/13 and 8/4, it diminishes with contrast and in unit 22/16, it increases with contrast. But in no case is the change in A more than a factor of 2 over the eightfold range of contrasts.

High-pass stage. The major effect of contrast is on the values of the parameters τ_H and k, which define the high-pass component of the form (1). The individual behaviour of these parameters is somewhat erratic; they both increase with contrast in unit 8/3, they both decrease in units 22/16 and 8/4, and they both show a maximum in unit 22/13.

One reason for this variability is that the time constant τ_H is so large; the high-pass component of eqn. (1) assumes a simpler form if $2f\tau_H$ is much greater than unity:

$$\frac{1}{1+[k/(1+2\pi if\tau_H)]} \rightarrow \frac{1}{1+[k/(2\pi if\tau_H)]}. \tag{2}$$

In this limit, which is approached for most of the data in Table 1 over most of the frequency range used, only the gain-bandwidth product, k/τ_H, is meaningful. Therefore, we have used that combination of parameters as a descriptor of the high-pass stage.

As can be seen in Table 1, the gain-bandwidth product k/τ_H always increases markedly and monotonically with contrast. The size of the increase varies from a factor of 2·5 in unit 22/13, to a factor of 7 in unit 8/3. Thus, the transductions we have measured become more high-pass with increasing contrast as if they were reflecting an increase in the gain-bandwidth product of an internal negative feed-back pathway.

X and Y cells. An interesting comparison can be made between the behaviour with contrast of the X and Y cells in Table 1. The values of the total delay $N_L \tau_L$ are approximately the same across cells; this indicates that similar low-pass filters feed

TABLE 1. The parameters of a model transfer function (eqn. (1)) that provide the best fit with observed first-order frequency responses from four ganglion cells. The first column of the table described each unit and the spatial frequency of the grating stimulus. In each case, the grating contrast was modulated by a sum-of-sinusoids signal at several levels of contrast per sinusoid, noted C. The model transfer function has an over-all gain of A. The low-pass component consists of N_L stages with time constant τ_L. The high-pass component consists of N_H stages, each of which has a low-pass filter of time constant τ_H in a feed-back loop with feed-back strength k. The values N_L and N_H listed in the first column are a compromise of the values N_L^{opt} and N_H^{opt} that were optimum at each contrast individually. R is the weighted squared deviation of the observed values and the fitted values with uniform N_L and N_H. $R-R^{opt}$ is the penalty in increased squared deviation incurred by forcing N_L and N_H to be uniform within each unit

Unit	C	A (impulses/sec)	τ_L (msec)	τ_H (sec)	k	$N_L\tau_L$ (msec)	k/τ_H (sec^{-1})	R	N_L^{opt}	N_H^{opt}	$R-R^{opt}$
8/3, Y on	0·0125	538	3·39	0·50	5·4	40·7	10·74	0·125	11	1	0·001
0·25 cycles/deg	0·025	573	3·65	0·56	7·6	43·9	13·7	0·028	11	1	0·003
$N_L=12$, $N_H=1$	0·05	441	3·40	0·70	13·8	40·8	19·6	0·049	13	1	0·000
	0·10	537	3·50	0·82	60·6	42·0	73·7	0·038	16	1	0·009
22/13, Y on	0·0125	277	3·52	1·26	8·9	44·2	7·1	0·165	8	1	0·020
0·4 cycles/deg	0·025	281	3·73	1·41	12·8	44·7	9·1	0·045	10	1	0·007
$N_L=12$, $N_H=1$	0·05	192	3·41	0·88	8·8	40·9	10·0	0·016	11	1	0·001
	0·10	156	3·18	0·31	5·9	38·2	19·2	0·047	17	1	0·016
8/4, X off	0·0125	412	1·98	1·37	7·8	47·6	5·7	0·062	24	1	0·000
0·25 cycles/deg	0·025	367	1·93	0·73	6·6	46·4	9·1	0·066	27	1	0·001
$N_L=24$, $N_H=1$	0·05	339	1·81	0·45	5·1	43·4	11·5	0·083	26	1	0·001
	0·10	254	1·75	0·15	3·1	42·1	20·70	0·082	20	1	0·002
22/16, X on	0·025	145	2·39	1·29	4·4	38·2	3·4	0·025	24	1	0·014
0·25 cycles/deg	0·05	153	2·42	0·69	3·8	38·8	5·6	0·024	5	1	0·000
$N_L=16$, $N_H=1$	0·10	190	2·52	0·15	3·2	40·4	21·8	0·050	14	1	0·001

into both types of cell. Of the three cells tested at 0·25 cycles/deg the Y cell reached the largest value of the feed-back gain-bandwidth product, k/τ_H. At the highest contrast (0·1) the value of k/τ_H for the Y cell was about 74 while the values for the two X cells were about 21 and 22. At 0·025 contrast the value of k/τ_H for the Y cell was only 13·7 while for the X cells it was 9·1 and 3·4. This is consistent with previous results which implied that the dynamics of Y cells are more different from X cells at high contrast (Shapley & Victor, 1978).

Spatial frequency. Next we consider the combined effects of spatial frequency and contrast on ganglion cell responses. The contributions of centre-surround antagonism may be separated from those of the contrast gain control by studying the spatial frequency dependence of model parameters at low contrast. Then to assess the contrast gain control's action one analyses the contrast dependence across spatial

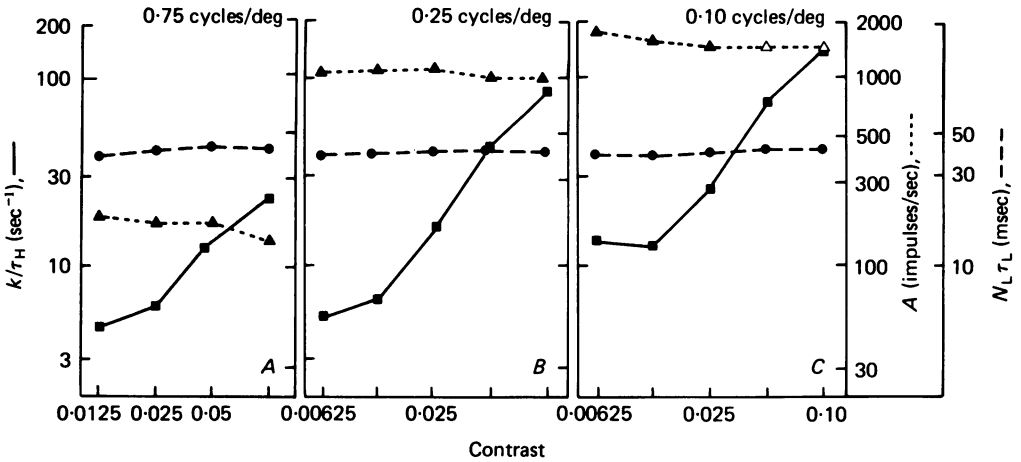

Fig. 3. The dependence of the model parameters on spatial frequency and contrast. A grating was used to evoke a first-order response in an on-centre Y cell. The parameters of overall gain, A, and low-pass delay, $N_L\tau_L$, are essentially independent of contrast ($N_L = 15$ in all cases). The high-pass gain-bandwidth product, k/τ_H, depended strongly on contrast. The values of A indicated by open triangles were not independent estimates because as k/τ_H grew large, the functional form of eqn. (1) became proportional to $A\tau_H/k$. Unit 34/3.

frequency. It should be understood by the reader that we do not imply by this formal analysis that centre-surround interaction proceeds via modulation of an internal negative feed-back. Abundant evidence exists to support the idea that centre-surround interaction in cat retinal ganglion cells may be thought of as a subtraction of the independent responses of two more or less independent response mechanisms, the centre mechanism and the surround mechanism (Rodieck, 1965; Enroth-Cugell & Robson, 1966; Enroth-Cugell & Pinto, 1972). All we aim to show in this section is that there is a systematic shift in the first-order frequency response of ganglion cells with spatial frequency and contrast which can be quantified and distinguished in terms of our model.

This programme to separate the effects of centre–surround antagonism from those of the contrast gain control is pursued in the analysis of the results illustrated in Fig. 3. The first-order frequency response of an on-centre Y cell was measured over a range

of contrast and spatial frequencies. The behaviour of the over-all gain A, the low-pass delay $N_L \tau_L$, and the high-pass gain-bandwidth product k/τ_H were determined by the least-squares fitting procedure described above. The low-pass delay $N_L \tau_L$ was about 40 msec at all spatial frequencies. The overall gain A increased monotonically as spatial frequency decreased.

TABLE 2. Spatial frequency and contrast dependence of k/τ_H in an X cell. This Table contains the derived parameter k/τ_H from the model of eqn. 1. The fit was performed with the number of stages of temporal integration, N_L, equal to 18

Spatial frequency (cycles/deg)	Contrast per sinusoid			
	0·0125	0·025	0·05	0·10
0·25	3·88	5·67	45·7	97·0
0·5	3·37	4·29	8·78	19·2
1·0	2·08	2·64	3·60	5·85
1·5	1·47	1·37	1·51	2·51

The value of the gain-bandwidth product in the absence of a contrast signal can be estimated by extrapolating from its values at the lowest contrasts. The value of k/τ_H in the absence of a contrast signal at 0·1 cycle/deg was approximately 13, while at 0·25 and 0·75 cycles/deg the low-contrast value of k/τ_H was at most 5. The increase in low-contrast k/τ_H as spatial frequency decreased is probably a manifestation of linear centre–surround interaction, independent of the contrast gain control. Thus, in this unit, centre–surround antagonism became a significant factor in determining response dynamics at spatial frequencies below 0·25 cycles/deg.

A similar dependence on spatial frequency and contrast was observed in X cells. Results of the analysis of data from a representative X cell are given in Table 2. By examining the values of the gain-bandwidth product k/τ_H, one can see that there is a gradual monotonic decrease in k/τ_H as spatial frequency increases. For each spatial frequency, the value of k/τ_H is always least at the lowest contrast. At the lowest contrast (0·0125 per sinusoid), the increase in the ratio k/τ_H may be taken as an indicator of centre–surround interaction. At the lower spatial frequencies the centre and surround of the receptive field are both stimulated and their interaction produces an increase in the k/τ_H required to fit the data. However, the systematic increase in k/τ_H with contrast at all spatial frequencies is the result of the action of the contrast gain control. As can be seen from the Table, the effect on k/τ_H produced by contrast is far greater than the effect produced by centre–surround interaction.

The data of Fig. 3 and Table 2 show that two fundamentally different processes influence the shape of the first-order frequency response: linear centre–surround antagonism and the nonlinear contrast gain control. Because the strength of the contrast gain control depends on spatial frequency, measurements of response dynamics at high contrast show a different spatial dependence from that indicated by measurements of response dynamics at low contrasts.

Quantitative comparison of spatial summation experiments and grating responses. The notion of a contrast gain control implies that one should also observe certain nonlinearities of spatial summation (Shapley & Victor, 1979). The spatial phase insensitivity of the contrast effect suggests that the contrast gain control sums contrast over a wide region of space (Shapley & Victor, 1978). Therefore, the response

of a ganglion cell to modulation in one region should be altered by the presence of a large nearby region of modulated light. This prediction has been verified (Shapley & Victor, 1979); here we analyse this spatial non-linearity in terms of our model for the first-order retinal pathway.

The spatial summation experiments to be considered consisted of measurements of the first-order frequency responses of a ganglion cell to four stimuli. $K_{1;C}(F)$ is the first-order response to a small spot positioned over the receptive field centre; $K_{1;P}(F)$ is the first-order response to synchronous modulation of four large regions in the receptive field periphery; $K_{1;C+P}(F)$ is the observed first-order response to combined central and peripheral modulation, and $K_{1;C-P}(F)$ is the observed first-order response to centre and periphery modulated in antiphase. The hypothesis of purely additive spatial combination of neural signals predicts that

$$\left.\begin{aligned} K_{1;C+P}(F) &= K_{1;C}(F)+K_{1;P}(F) \\ \text{and} \quad K_{1;C-P}(F) &= K_{1;C}(F)-K_{1;P}(F). \end{aligned}\right\} \tag{3}$$

Systematic departures from the simple additive prediction (eqn. 3) were observed in all ganglion cells (X and Y) studied (Shapley & Victor, 1979). These departures were analysed by extracting two 'effective frequency responses':

$$\left.\begin{aligned} K'_{1;C}(F) &= \tfrac{1}{2}(K_{1;C+P}(F)+K_{1;C-P}(F)) \\ \text{and} \quad K'_{1;P}(F) &= \tfrac{1}{2}(K_{1;C+P}(F)-K_{1;C-P}(F)) \end{aligned}\right\} \tag{4}$$

The effective frequency response $K'_{1;C}$ and $K'_{1;P}$ are the unique quantities whose algebraic sum and difference yield the observed combined responses $K_{1;C+P}$ and $K_{1;C-P}$. Since the response of each region is altered by the presence of modulation in the other region, the effective frequency response may be considered to be the response of one region in the presence of modulation in the other region.

In the experiments to be analysed here, we used a spatial configuration consisting of a small central spot and large peripheral regions. Under these conditions, the effective periphery response $K'_{1;P}$ was nearly identical to that of the periphery response in isolation, $K_{1;P}$ (Shapley & Victor, 1979). Thus, the small central spot had little effect on the periphery response. However, the effective centre response $K'_{1;C}$ was always attenuated at low frequencies, and phase-advanced at high frequencies, with respect to the isolated centre response, $K_{1;C}$ (Shapley & Victor, 1979). This effect was present even when the peripheral stimulus was a modulated grating which produced no response by itself.

Next we consider an analysis of isolated centre responses and effective centre responses in terms of the model transfer function (1). The results of this analysis for an X cell and a Y cell are presented in Table 3. The spatial dimensions of the central and peripheral stimuli are given in the table. For comparison, Table 3 also gives the parameters obtained for responses of these units to gratings at a series of contrasts. The basic procedure described above was used to obtain the parameters A, N_L, τ_L, τ_H, and k of eqn. (1). The effect of peripheral stimulation on the centre's response can be seen by comparing the effective centre's response with the isolated centre's response. The total delay of the low-pass stages, $N_L\,\tau_L$, is unchanged by the presence

TABLE 3. A comparison of the effect of contrast and non-linear spatial summation on the parameters of the model transfer function. The parameters A, N_L, τ_L, N_H, τ_H, and k determine the overall gain, the low-pass stages, and the high-pass stages of the model transfer function (eqn. (1)), as described in the legend of Table 1. R is the weighted squared deviation of the observed values of the first-order responses and the fitted values. Each of the two units was tested under two conditions; a full-field grating at a series of contrasts, and a summation experiment in the 'flag' configuration (Shapley & Victor, 1979*b*). For the full-field grating experiments, the contrast per sinusoid is denoted C. For summation experiments, the isolated centre responses and the effective centre responses were fit with the empirical transfer function (eqn. (1)). The conditions for the summation experiments were: unit 18/4 – centre spot, 0·8 × 1·0 deg; separation of peripheral regions from centre spot, 1·8 deg: unit 17/8 – centre spot, 1·2 × 1·2 deg; separation of peripheral regions from centre spot, 1·5 deg. In both units, the summation experiment was also performed using a grating as the peripheral stimulus; the contrast of the grating was twice that of the uniform stimuli, to provide an equal spatial root-mean-square contrast

Unit	Configuration		A (impulses/sec)	$N_L \tau_L$ (msec)	k/τ_H (sec^{-1})	R
18/4 X on	Grating experiment	$C = 0{\cdot}0125$	416	44·5	2·7	0·042
		$C = 0{\cdot}025$	421	44·5	4·3	0·053
	0·7 cycles/deg	$C = 0{\cdot}05$	372	44·7	8·4	0·030
		$C = 0{\cdot}10$	306	45·3	14·5	0·029
$N_L = 22$						
$N_H = 1$						
	Summation experiment $C = 0{\cdot}025$					
	auxiliary peripheral stimulus:					
	uniform	isolated centre	533	50·5	2·4	0·044
		effective centre	652	48·6	8·5	0·017
	0·7 cycles/deg	isolated centre	526	47·4	2·4	0·044
		effective centre	614	46·2	7·2	0·049
17/8 Y on	Grating experiment	$C = 0{\cdot}0125$	461	35·1	2·9	0·043
		$C = 0{\cdot}025$	466	33·2	4·8	0·025
	0·25 cycles/deg	$C = 0{\cdot}05$	398	35·4	6·9	0·035
		$C = 0{\cdot}10$	277	36·2	9·9	0·033
$N_L = 13$						
$N_H = 1$						
	Summation experiment $C = 0{\cdot}05$					
	auxiliary peripheral stimulus:					
	uniform	isolated centre	355	39·9	4·5	0·157
		effective centre	403	42·4	18·3	0·107
	1·0 cycles/deg	isolated centre	338	39·1	4·7	0·059
		effective centre	358	36·5	10·0	0·019

of peripheral stimulation. However, the gain-bandwidth product for the high-pass stage, k/τ_H, increases by a factor of 2·5–4 in the presence of peripheral stimulation. This is the combination of parameters that shows a similar marked increase with contrast in the uniform grating experiments. The increase in k/τ_H is somewhat greater for a uniform peripheral stimulus than for an auxiliary peripheral stimulus consisting of a fine grating. However, peripheral grating stimuli which elicited no response from the X cell and only a second-order response from the Y cell still produced an increase of k/τ_H by a factor of 2·5–3, as can be seen in Table 3.

The comparison of the effect of increasing the contrast of a grating with the effect of an auxiliary stimulus on the centre's response can also be made in a parameter-free manner. The effect of increasing the contrast of a grating stimulus may be expressed

as the ratio of the frequency responses at two contrasts, normalized by the ratio of the two contrasts. The effect of an auxiliary peripheral stimulus on the centre response may be expressed as the ratio of the effective centre response to the isolated centre response. In both cases, the ratio is a complex number. The amplitude of the complex number is the relative amplitude change induced by the experimental manipulation, and the phase of the complex number is the phase change induced by the manipulation. Fig. 4 shows the results of this calculation for the Y cell of Table 3. Both the amplitude and phase changes induced by increasing the contrast of a grating stimulus correspond closely to the changes induced by adding an auxiliary peripheral stimulus. This further supports the identification of the mechanism responsible for non-linear spatial summation with the contrast gain control.

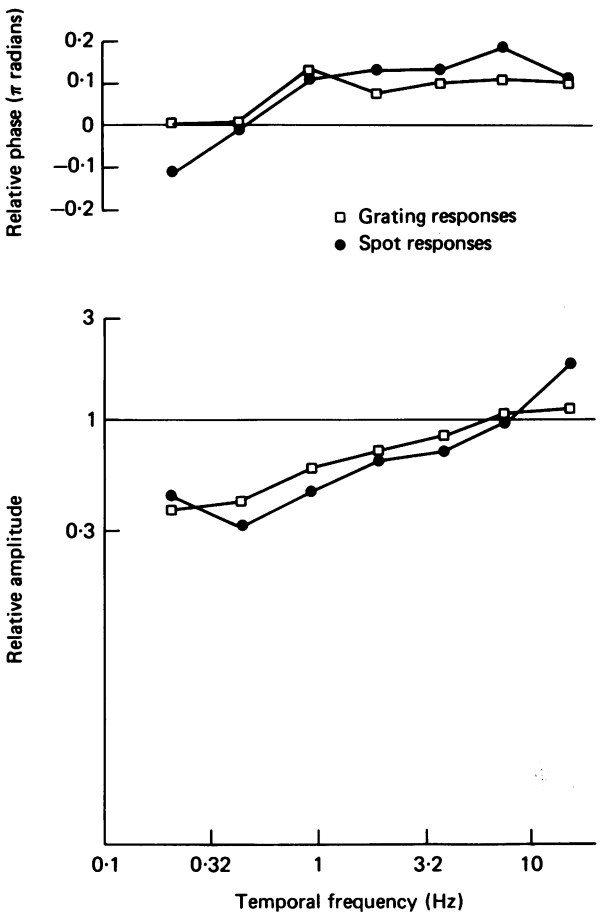

Fig. 4. A comparison of the effects of increased contrast and of an auxiliary stimulus. The response of a Y cell to an 0·25 cycles/deg grating positioned to produce a maximal first-order response was measured at contrasts of 0·025/sinusoid and 0·1/sinusoid. The ratio of these responses, divided by 4 (the ratio of the input contrasts), were plotted as open squares. The ratio of effective centre responses to isolated centre response in a summation experiment performed on this cell were plotted as filled circles. These ratios, which express the change in shape of the first-order response as a result of the two experimental manipulations, are seen to be similar functions of temporal frequency. Unit 17/8.

The effect of contrast on periphery responses. The effect of peripheral stimulation on the centre's response is not a consequence of centre–surround interaction *per se*, because a grating pattern in the periphery which does not stimulate the classical linear surround generates a strong effect on the centre response (Table 3). The hypothesis of spatial spread of the contrast signal explains this result, and this hypothesis also explains why the central stimulus does not alter the surround response. The central spot is sufficiently small in area so as not to raise the contrast signal appreciably above the value set by the peripheral stimulus itself. This reasoning suggests that a dependence of the surround response on the presence of auxiliary stimulation could be demonstrated, provided that the auxiliary stimulus was large enough in spatial extent. The next results demonstrate this predicted effect of contrast and auxiliary stimuli on test stimuli located in the receptive field periphery.

Fig. 5 shows the response of an off-centre Y cell to a 6·5 × 20 deg bar in the receptive field periphery, modulated by a sum-of-sinusoids signal at two levels of contrast. The first-order responses illustrated showed the same kind of qualitative dependence on contrast as did the grating responses (Fig. 1). As contrast increased over an eightfold range, the low-frequency responses increased in amplitude by a factor of 3, but the high-frequency responses increased in amplitude by a factor of about 10. Responses in the range 1–15 Hz showed a phase advance of 0·2–0·3 π radians over this contrast range.

The model transduction (1) was also fit to responses from the receptive field periphery. Table 4 shows the dependence of the parameters of the model on input contrast, for both the Y cell of Fig. 5 and an X cell. The fitted frequency responses are compared with the observed responses in Fig. 5 to demonstrate a reasonably good fit of the model with data.

Inspection of Table 4 shows that the dependence of the model parameters on contrast is very similar to the dependence observed in the grating experiments (Table 1). The overall gain, A, and the low-pass delay, $N_L \tau_L$, are essentially independent of input contrast. As in the dependence of grating-driven responses on contrast, only the gain-bandwidth product for the high-pass stage, k/τ_H, was affected systematically by increasing the stimulus contrast. The increase in the gain-bandwidth product was substantial, a factor or more than six over a contrast range of eight. The form of the dependence of periphery responses on contrast is thus similar to that of the grating-driven responses. However, the contrast-independent low-pass component of the periphery transduction is fit best by a series of more low-pass stages (N_L = 20–80) with greater total delay ($N_L \tau_L$ = 45–60 msec), than the low-pass component of responses that contain a contribution from the centre mechanism for which N_L = 12–24 and $N_L \tau_L$ = 35–45 msec. This difference probably reflects additional stages of temporal integration or delay in the generation of surround antagonism.

The effect of auxiliary stimulation on periphery responses. Responses from the receptive field periphery provide an opportunity to compare directly the effect of increased stimulus contrast with the effect of an auxiliary stimulus. We examined the combination of responses from the receptive field periphery by using stimuli consisting of two bars on opposite sides of the receptive field centre. The two peripheral stimuli in isolation gave responses denoted by $K_{1;P_1}(F)$ and $K_{1;P_2}(F)$. The responses to the two regions modulated in synchrony was called $K_{1;P_1+P_2}(F)$, and the

response to the two regions modulated in antiphase was called $K_{1;P_1-P_2}(F)$. As in the centre/periphery experiments, the ganglion cell responses to the combined stimuli deviated consistently from the prediction of simple additivity (Victor, 1979). This deviation was analysed in terms of the effective periphery responses (cf. eqn. 4) defined by

$$K'_{1;P_1}(F) = \tfrac{1}{2}(K_{1;P_1+P_2}(F) + K_{1;P_1-P_2}(F)) \quad \text{and} \quad K'_{1;P_2}(F) = \tfrac{1}{2}(K_{1;P_1+P_2}(F) - K_{1;P_1-P_2}(F)) \tag{5}$$

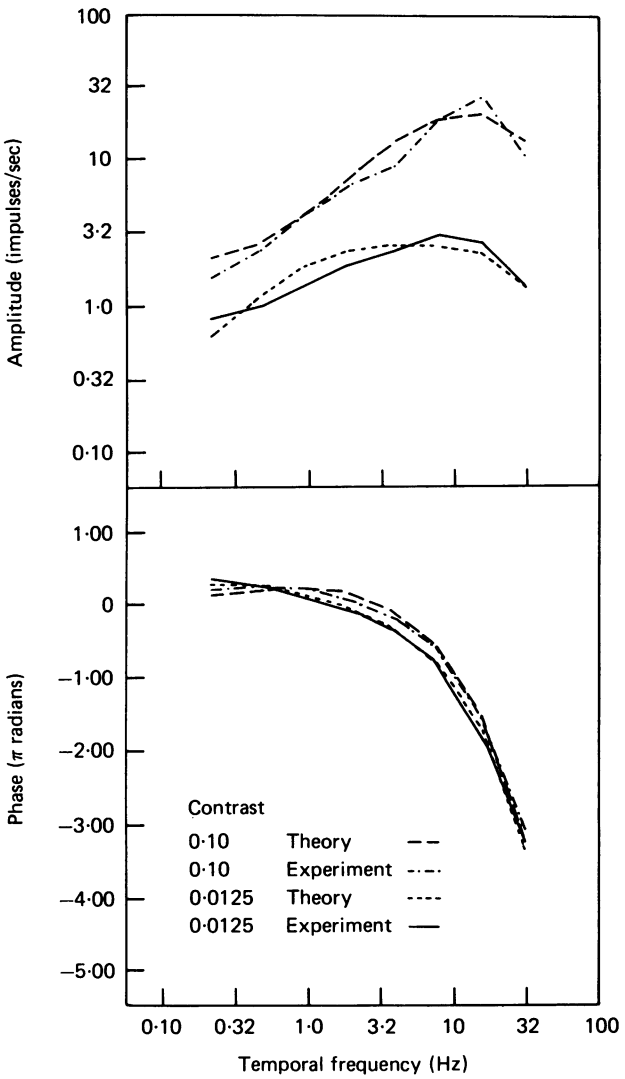

Fig. 5. A comparison of measured first-order responses to a peripheral bar at high and low contrasts with the response of a lumped model. An off-centre Y cell was stimulated with a 6·5 × 20 deg bar in the receptive field periphery. The eight-sinusoid sum produced a contrast per sinusoid of 0·0125 (——) and 01·1 (–·–·). The first-order frequency responses of a model transducer (eqn. (1)) are shown for contrast levels of 0·0125 (. . .) and 0·1 (– – –) per sinusoid. The parameters of the model are given in Table 3. Unit 42/7.

Consider the idea that the deviation from simple additive spatial summation is due to spatial spread of the contrast signal. From previous results, we had concluded that the contrast gain control summed contrast signals over a wide retinal area. If one assumes that its sensitivity is approximately constant with position, one may predict that the effect of increased area of stimulation would be equivalent to increased average contrast, provided that the additional area of stimulation is sufficiently nearby to the test region to allow effective spread of the contrast signal. Therefore we expect that the effective response to a test bar in the presence of a second bar of the same area and contrast should be similar to the response to the test bar at twice the contrast. This prediction is confirmed by the data in Fig. 6. The response

TABLE 4. The parameters of model transfer functions (eqn. (1)) that provide the best fit with observed first-order responses evoked by peripheral stimuli. The meaning of the parameters is described in the text and the legend of Table 1. The entry marked * is a blank, because, as explained in the legend of Fig. 3, A could not be estimated independently. Increased contrast per sinusoid, C, results in a selective increase in the gain-bandwidth product of the high-pass stage, k/τ_H

Unit	C	A (impulses/sec)	$N_L \tau_L$ (msec)	k/τ_H (sec^{-1})	R
42/7, Y off $N_L = 80, N_H = 1$					
isolated periphery 6·5 × 20 deg	0·0125	217	54·7	5·72	0·086
	0·025	236	53·8	7·80	0·085
	0·05	290	52·9	23·33	0·097
	0·10	249	52·1	36·73	0·089
35/1, X on $N_L = 26, N_H = 1$					
isolated periphery 7·5 × 20 deg	0·0125	738	46·9	4·60	0·046
	0·025	717	46·5	7·71	0·033
	0·05	—*	48·6	28·95	0·030

to a 7·5 × 20 deg bar, $K_{P_1}(F)$, was obtained at three levels of contrast, and the parameters $N_L \tau_L$ and k/τ_H were extracted as described above. At each contrast level, the effective response $K'_{P_1}(F)$ in the presence of a second bar placed on the opposite side of the centre was determined according to eqn. (5). The effective responses were similar to the isolated responses at twice the contrast. The parameter k/τ_H, which is selectively influenced by the contrast signal, assumed nearly the same value in episodes equated for the product contrast times area. Thus, the effectiveness of a visual stimulus in activating the contrast gain control is a function of the product of the stimulus contrast and area.

DISCUSSION

The present results, which demonstrate a fractionation of the first-order frequency response into contrast-dependent and contrast-independent components, give strong support to our previous inferences which were made initially on the basis of a qualitative comparison of curves (Shapley & Victor, 1978, 1979). The effect of increased contrast was nearly identical to the effect of a nearby auxiliary stimulus (Fig. 4; Table 2), supporting the notion that the spatial spread of the contrast signal

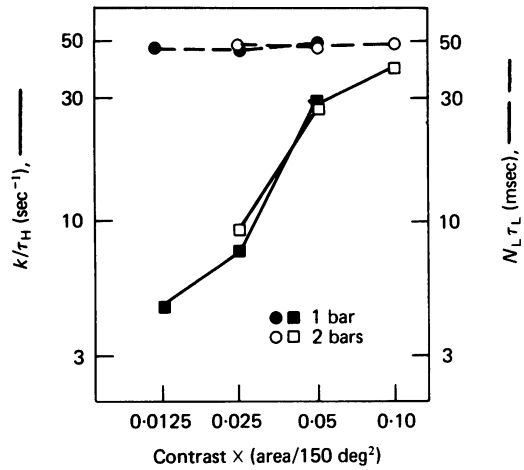

Fig. 6. The dependence of the first-order response on contrast and area of stimulation in the periphery of an on-centre X cell. The parameters $N_L\,\tau_L$ and k/τ_H describe the low-pass and high-pass components of a model transfer function (eqn. (1)) that was fit to the data obtained with a stimulus consisting of a single 7·5 × 20 deg bar in the receptive field periphery. Open symbols are derived from data obtained with a stimulus consisting of two similar bars on opposite sides of the receptive field centre, separated by 1·5 deg. It is seen that both data sets show the same dependence on the contrast-area product. Unit 35/1.

is responsible for the non-linearity of spatial summation. This point is also supported by the equivalent effects of adding an auxiliary stimulus of equal area, or of doubling the contrast of the original stimulus, in experiments on the receptive field surround (Fig. 6).

The model transfer function (eqn. 1) helps to restrict the possibilities for the site of action of the contrast signal. The low-pass stages of the centre transduction are similar in all units studied (both X and Y), and independent of contrast. This portion of the transduction is probably associated with the photoreceptors. Thus, the contrast signal acts at a stage of retinal processing subsequent to the photoreceptors. On the other hand, a qualitative analysis of the effect of contrast on the second-order response of Y cells (Shapley & Victor, 1980) showed that the contrast signal acts primarily before the generation of the nonlinearity, which we believe takes place in the amacrine cell layer. Thus, it seems that the contrast gain control represents a non-linear feed-back from the subunit pathway, either entirely within the inner plexiform layer, or via a neurone such as the interplexiform cell (Boycott, Dowling, Fisher, Kolb & Laties, 1975) to the outer plexiform layer.

The model reported here is in some ways similar to but in other ways different from previous attempts to provide a theory for non-linearities in the function of the retina. The functional form for the low-pass transductions is similar to the filter cascade models of de Lange (1952) and Fuortes & Hodgkin (1964). Based on the curve fitting of our empirical results, we find that the time constants of these low-pass filters, which act as leaky integrators, are independent of contrast (or depth of modulation) over the range of stimuli which we have used. This finding distinguishes our model from the Fuortes–Hodgkin model in which the time constants of the filters in the cascade are changed by an amount which depends on the output of the last stage in the

cascade. However, there is another important way in which our model of the change in frequency response with contrast differs from the Fuortes–Hodgkin model and other proposed models of retinal nonlinearities (for a review of several other such models, cf. van de Grind, Grüsser & Lunkenheimer, 1973). We propose that the non-linear subunit pathway, which is parallel to the direct route of signal transmission from photoreceptors to ganglion cells, provides the non-linear excitatory responses of Y cells and also feeds back, in the way described above, to modify signal transmission along the direct route. The spatially distributed nature of the non-linear feed-back signal in our model distinguishes it from the locally non-linear models of, for example, Fuortes & Hodgkin (1964) and DeVoe (1967).

Previous models which dealt with non-linear spatial interaction were based on the original paper of Furman (1965) on shunting inhibition. A particularly explicit model of this sort was offered by van de Grind *et al.* (1973). Such shunting inhibition models only account for non-linearities of amplitude; they are essentially static models. However, in our experimental results amplitude non-linearities and phase non-linearities were always inextricably linked. Therefore, an adequate model must include a mechanism which affects amplitude and phase simultaneously. This is the constraint which led to the formulation of the model in eqn. (1).

A discussion of the lumped model

The intention of formulating the lumped model (eqn. 1) was to provide a quantitative way of describing the shape of the measured first-order frequency kernels. Thus, in constructing a model transfer function, we found it to be more useful to keep the number of parameters small than to provide a term for all of the possible stages of retinal processing. As a consequence, only one kind of low-pass stage, with a single time-constant, was used. This may be one of the reasons that the optimum number of low-pass stages, N_L, was usually in the range of 12–24. This number is large in comparison to the number of low-pass stages required to fit the transfer functions of vertebrate photoreceptors (Pasino & Marchiafava, 1976), and thus may reflect the presence of additional low-pass stages proximal to the receptors. Another possibility is that the present data extend to higher temporal frequencies than the photoreceptor data. A six-stage low-pass filter, which fits photoreceptor data well, would have a phase shift that reaches a high-frequency asymptote of 3π radians, and would fit the low-frequency portion of the present data well. The high-frequency portion of our data (Fig. 1) clearly shows that phase continues to fall rapidly.

Another component of the retinal transduction contributing phase lag at very high temporal frequencies is conduction delay in the ganglion cell axon. The conduction time of the nerve impulse from the ganglion cell body to the recording site in the optic tract is about 2 msec for Y cells and 4 msec for X cells (Stone & Fukuda, 1974). The upper figure for the delay would result in a phase shift of about $0{\cdot}25\,\pi$ radians at the highest temporal frequency used. By inserting a pure delay factor $e^{-i2\pi f\delta}$ in eqn. (1), one can obtain a good fit with fewer low-pass stages. However, since the conduction delay was not determined experimentally, this embellishment would introduce another free parameter into the model without yielding additional insight. Therefore, it was not added to the model. A further contribution to the low-pass transductions we have measured may come from intraretinal but post-receptoral delays. Such

processes have been deduced from horizontal cell transfer functions (Foerster, van de Grind & Grüsser, 1977). If such delays do contribute, our data imply that they are unaffected by contrast.

This work was supported by grants from the U.S. National Eye Institute EY 188, EY 1428, and EY 1472. R.S. was also supported by a Career Development Award from the Eye Institute.

REFERENCES

BAYLOR, D. A., HODGKIN, A. L. & LAMB, T. D. (1974). Reconstruction of the electrical responses of turtle cones to flashes and steps of light. *J. Physiol.* **242**, 759–791.

BOYCOTT, B. B., DOWLING, J. E., FISHER, S. K., KOLB, H. & LATIES, A. M. (1975). Interplexiform cells of the mammalian retina and their comparison with catecholamine-containing retinal cells. *Proc. R. Soc.* B **191**, 353–368.

DE LANGE, H. (1952). Experiments on flicker and some calculations on an electrical analogue of the foveal systems. *Physica* **18**, 935–950.

DEVOE, R. D. (1967). A nonlinear model for transient responses from light-adapted wolf spider eyes. *J. gen. Physiol.* **50**,1993–2030.

DOWLING, J. E. & BOYCOTT, B. B. (1966). Organization of the primate retina: Electron microscopy. *Proc. R. Soc.* B **166**, 80–111.

ENROTH-CUGELL, C. & PINTO, L. (1972). Properties of the surround response mechanisms of cat retinal ganglion cells and centre–surround interaction. *J. Physiol.* **220**, 403–439.

ENROTH-CUGELL, C. & ROBSON, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol.* **187**, 517–552.

FLETCHER, R. & POWELL, M. J. D. (1963). A rapid descent method for minimization. *Computer J.* **6**, 163–168.

FOERSTER, M. H., VAN DE GRIND, W. A. & GRÜSSER, O.-J. (1977). The response of cat horizontal cells to flicker stimuli of different area, intensity and frequency. *Exp. Brain Res.* **29**, 367–385.

FUORTES, M. G. F. & HODGKIN, A. L. (1964). Changes in time scale and sensitivity in the ommatidia of *Limulus*. *J. Physiol.* **172**, 239–263.

FURMAN, G. G. (1965). Comparison of models for subtractive and shunting lateral inhibition in receptor neuron fields. *Kybernetik* **2**, 257–274.

HOCHSTEIN, V. & SHAPLEY, R. M. (1976). Quantitative analysis of retinal ganglion cell classifications. *J. Physiol.* **262**, 237–264.

NAKA, K.-I., MARMARELIS, P. Z. & CHAN, R. (1975). Morphological and functional identification of catfish retinal neurons. III. Functional identification. *J. Neurophysiol.* **38**, 92–131.

PASINO, E. & MARCHIAFAVA, P. L. (1976). Transfer properties of rod and cone cells in the retina of the tiger salamander. *Vision Res.* **16**, 381–386.

RODIECK, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Res.* **5**, 583–601.

SHAPLEY, R. M. & ROSSETTO, X. (1976). An electronic visual stimulator. *Behav. Res. Meth. & Instr.* **8**, 15–20.

SHAPLEY, R. M. & VICTOR, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *J. Physiol.* **285**, 275–298.

SHAPLEY, R. M. & VICTOR, J. D. (1979). Non-linear spatial summation and the contrast gain control of cat retinal ganglion cells. *J. Physiol.* **290**, 141–161.

SHAPLEY, R. M. & VICTOR, J. D. (1980). The effect of contrast on the non-linear response of the Y cell. *J. Physiol.* **302**, 535–547.

STONE, J. & FUKUDA, Y. (1974). Properties of cat retinal ganglion cells: A comparison of W cells with X and Y cells. *J. Neurophysiol.* **37**, 722–748.

TOYODA, J.-I. (1974). Frequency characteristics of retinal neurons in the carp. *J. gen. Physiol.* **63**, 214–234.

VAN DE GRIND, GRÜSSER, O. J. & LUNKENHEIMER, H. V. (1973). Temporal transfer properties of the afferent visual system. In *Handbook of Sensory Physiology* VII/3A, Central Processing of Visual Information A: Integrative Functions and Comparative Data, ed. JUNG, R., pp. 431–524. Berlin, Heidelberg and New York: Springer-Verlag.

VICTOR, J. D. (1979). The functional organization of the receptive fields of cat X and Y retinal ganglion cells. Ph.D. Thesis, Rockefeller University, New York.

VICTOR, J. D. & SHAPLEY, R. M. (1979). Receptive field mechanisms of cat X and Y retinal ganglion cells. *J. gen. Physiol.* **74**, 275–298.

VICTOR, J. D. & SHAPLEY, R. M. (1980). A method of nonlinear analysis in the frequency domain. *Biophys. J.* **29**, 459–484.