QUANTITATIVE ANALYSIS OF RETINAL GANGLION CELL CLASSIFICATIONS

BY S. HOCHSTEIN* AND R. M. SHAPLEY

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

(Received 16 December 1975)

SUMMARY

1. The classification of cat retinal ganglion cells as X or Y on the basis of linearity or nonlinearity of spatial summation has been confirmed and extended. Recordings were taken from optic tract fibres of anaesthetized, paralysed cats.

2. When an alternating phase sine wave grating was used as a stimulus, X cells had null positions and Y cells responded at all positions of the grating.

3. These results did not depend on the temporal wave form or the temporal frequency of pattern alternation over a wide range.

4. At high spatial frequencies for the particular cell, a Y cell gave a big 'on-off' response, or frequency doubling, at all positions of the grating, while an X cell did not.

5. The use of contrast sensitivity versus spatial phase also served to differentiate the two cell types. With an alternating sine grating stimulus X cells had a sinusoidal dependence on spatial phase, while each Y cell's sensitivity depended in a complicated manner on spatial phase.

6. Sensitivity versus spatial phase for different Fourier components of the neural response also separated the two classes of cells. Significant second harmonic distortion was present in Y cells. The second harmonic component was spatial phase insensitive, and became dominant at high spatial frequencies.

7. The maximum of the 2nd/1st harmonic ratio was taken as an index of nonlinearity. X cells always had a nonlinearity index less than 1 while in Y cells this index always exceeded 1.

8. Response to spots, diffuse light and drifting gratings were compared to the nonlinearity index as a basis for classifying cells. The nonlinearity index was most reliable because it was least dependent on retinal eccentricity.

* Present address: Neurobiology Unit, Institute of Life Sciences, Hebrew University, Jerusalem, Israel.
INTRODUCTION

Knowledge about the physiology of retinal ganglion cells is useful in understanding the whole visual pathway; the axons of these cells are the only links between the retina and the rest of the brain. In measuring the impulse discharge of ganglion cells in response to spatial patterns, we are allowed to view the operation of the retinal network as the brain must view it. Two main categories of problems concern investigators in this field: (1) how to account for the responses of ganglion cells in terms of retinal wiring and (2) how responses of ganglion cells influence responses of more central neurones and hence visual perception.

This paper is about the process of spatial summation in retinal ganglion cells of the cat. On the foundation laid by Kuffler (1953), Rodieck (1965) had previously built an elegant linear theory of the operations which the retinal inputs to these cells perform on any spatio-temporal pattern. According to this view, all retinal ganglion cells summed excitation and inhibition from their receptive field centres and surround, weighted by the spatial sensitivities of the centre and surround mechanisms and the time course(s) of these mechanisms’ responses. One flaw of this simple conception of ganglion cell physiology was that it did not jibe with what was known about the complexity of retinal morphology (Cajal, 1892; cf. Rodieck, 1973; Shkolnik-Yarros, 1971).

The defect of simplicity has been more than remedied by the discovery of many types of cat retinal ganglion cells (Enroth-Cugell & Robson, 1966; Cleland, Dubin & Levick, 1971; Stone & Hoffmann, 1971; Ikeda & Wright, 1972; Cleland & Levick, 1974). If one looks in the right way, one can see now a close correspondence between the physiological and anatomical categories of ganglion cells (Boycott & Wässle, 1974; Levick, 1975). Enroth-Cugell & Robson (1966) initiated this era of categorization with their discovery of X and Y cells. They found that the X cells behaved like simple retinal ganglion cells as described by Kuffler (1953) and, more elaborately, by Rodieck & Stone (1965). Signals generated by light in photoreceptors added up linearly in X cells so that, for instance, zero response could be produced by the simultaneous presentation of an increment and decrement in different parts of the receptive field of the cell. In Y cells this simple cancellation of increment and decrement would not work. This indicated there was a non-linear stage before all the retinal input to the Y cell was pooled.

Many other tests have been used recently to distinguish between ganglion cell types in the cat retina (Cleland et al. 1971; Hoffmann, Stone & Sherman, 1972; Ikeda & Wright, 1972; Cleland, Levick & Sanderson, 1973) but there has been little or no attempt to correlate these other categoriza-
tions with the original X/Y classification based on linear/non-linear spatial summation. Our previous research on cells in the lateral geniculate nucleus (Shapley & Hochstein, 1975) led us to believe that the linear/non-linear summation property was an essential one, from which many of the other distinguishing characteristics were derived. The work reported in this paper is an attempt at a more rigorous linear/non-linear classification. It also is an attempted correlation of the linear/non-linear classification of X and Y cells with some of the other classification schemes people have used. The paper ends with conjectures about synaptic connexions in the cat retinal based on our results with X and Y cells.

**METHODS**

Cats were prepared for single unit recording by gluing a small cylindrical plastic cap over a craniotomy (0.25 in. in diameter), centred at stereotaxic co-ordinates A9L9 for work on the optic tract. Venous cannulae and a tracheal cannula were inserted. The e.k.g. and body temperature were measured and the latter was controlled by a heating pad. Anaesthesia was induced with ketamine and continued during surgery with sodium thiamylal (Surital), then maintained during the experiment with urethane (0.05 g/kg hr i.v. after a 0.2 g/kg i.v. loading dose). Muscle paralysis was accomplished by infusion of gallamine triethiodide (10 mg/kg hr i.v.) and diallyl-bis-nortoxiferine (0.25 mg/kg/hr i.v.), and the cat was artificially respired during the experiment. We also performed a bilateral cervical sympathectomy to reduce eye movements (Rodieck, Pettigrew, Bishop & Nikara, 1967). We gave the cats penicillin and dexamethasone to increase the longevity of the experiment which usually lasted 36 hr with the cat in good condition.

After the cat was paralysed and 10% phenylephrine hydrochloride and 1% atropine sulphate were applied to the eyes, the optic disks of both eyes were mapped on a tangent screen with an ophthalmoscope (Fernald & Chase, 1971). Contact lenses (+2D) with an artificial pupil 3 mm in diameter were then placed on the eye. Then we refracted the cat with the ophthalmoscope to check that it was in focus for a target 50 cm away. Periodically the physiological optics were checked and the eyes irrigated with saline. Occasionally the disks were mapped on the screen at the end of an experiment to ascertain if the eyes had moved significantly. They never were more than 2° away from the initial position.

The stimuli were produced on the screen of an oscilloscope by means of a set of electronic circuits which are described in detail elsewhere (Shapley & Rossetto, 1976). Patterns were created by the method of Schade (1956); a wave form synchronized to the sweep voltage of a TV type raster was displayed as a visual stimulus on the oscilloscope face by connecting it to the Z axis of the oscilloscope. The frame rate of the raster was 200 c/s. So a sinusoidal grating which had 10 cycles/screen width was created by synchronizing a 2 kC sine wave with the start of the raster sweep and leading the 2 kC to the Z axis input. The screen subtended 10° x 13° and was 57 cm away from the eye. At this distance therefore, a 2 kC sine wave input to the Z axis produced a visual sine grating with spatial frequency 10/13 c/deg. Linearity of the display was good up to 50% contrast.

Mean luminance was 1 cd/m². The mean luminance and contrasts were measured with a Spectra brightness spot meter (Photo Research Corp.). The oscilloscope has a P31 phosphor which appears to us as a desaturated yellow-green at this mean luminance.
The most useful stimulus pattern used in these experiments was the alternating phase, or contrast reversal, sine grating. This pattern was formed by multiplying the spatial signal, a high speed sine wave synchronized with the sweep, by a slow a.c. modulation voltage in an analog multiplier. The output of the multiplier then was connected to the Z axis of the oscilloscope. A PDF 11/20 computer was used to produce the slow modulation voltage (the machine read out the entries of a modulation list of numbers repetitively into a digital-to-analogue converter (DAC)). The computer also controlled the position of the alternating grating (by varying the delay of the gate of a gated oscillator (Tektronix FG501) with respect to the start of the sweep). Since the computer could control the position and also the depth of modulation at each position, it was able to perform a complete null test experiment automatically, and also was able to determine the contrast sensitivity at each position as described below. Also, one could programme the computer to use different temporal wave forms and various rates of temporal modulation as described below. A diagram of the alternating grating stimulus is given in Fig. 1.

Modulated aperiodic stimulus test patterns (diffuse light and bars) were formed by multiplying a proper width pulse, from a pulse generator, by the computer-generated temporal modulation signal. Spots were produced by leading the Z axis signal to a

Alternating phase grating

\[ L_1 + L_0 \]
\[ L_0 \]

0

On-off grating

\[ L_1 + L_0 \]
\[ L_0 \]

0

Fig. 1. Alternating phase (contrast reversal) grating and on-off grating stimuli. The top row of graphs is the luminance profile of an alternating phase grating at its two alternate phases one half cycle apart. The dotted vertical line indicates a fixed reference point on the stimulus screen. The bottom row of graphs is the luminance wave form of an on-off grating — also called introduction and withdrawal of a grating — with a spatial phase of 0° with respect to the top row of alternating grating. The mean luminance was \( L_0 \) and the contrast was \( L_1/L_0 \).
high speed switching circuit (a variable window circuit) which only allowed the signal for the pattern to go to the oscilloscope when the Y axis deflexion voltage was between prescribed limits. Thus we could control the length of a narrow bar, making it into a small square or rectangular spot. Comparison of our techniques with various standard categorization tests using these aperiodic spatial stimuli is made at the end of the Results section. Our versions of the standard tests of, for example, Cleland et al. (1971) were similar to those used in our measurements of linearity of spatial summation in that the mean luminance 1 cd/m², was the same, the contrasts used were in the range 0.075–0.3, and different temporal wave forms and frequencies were explored. Responses to drifting gratings at various spatial frequencies, drift rate, and contrasts were also measured; the drifting gratings were also produced on the oscilloscope face under the same range of stimulus conditions as the alternating phase gratings.

Single fibre recording was performed with tungsten-in-glass micro-electrodes (Levick, 1972). Amplified nerve impulses were led to an oscilloscope and audio-monitor. The electrode was advanced until a single impulse-form stood out above all others. A comparator circuit produced a fixed square-pulse for each spike recorded. This pulse was transmitted to the PDP 11/20 computer which binned the responses of each cycle into, typically, 60 bins and produced a PST (post-stimulus time) histogram by adding up the responses within each bin for all cycles. Total duration of each run was usually 15 sec (60 sec only when stimulus frequency was 0.05 Hz) with number of repetitions of the stimuli determined by the temporal frequency. The computer simultaneously produced a second histogram of number of spikes per sweep versus the number of the sweep in order to monitor slow trends like habituation. No strong habituation effects were found. By dividing the impulse/bin in the histogram by the total averaging time for that bin (bin duration times number of sweeps) we could compute the averaged instantaneous impulse rate in units of impulses/sec.

Spatial stimuli were always repeated at two or three different contrasts in order to determine contrast sensitivities (see Results). The paradigm was as follows: the first run with a particular stimulus was performed at an intermediate contrast level (0.15 contrast). Then the computer compared the averaged response with a predetermined criterion level. (The peak averaged firing rate minus the minimum firing rate was used as a first measure of the response amplitude.) If the response exceeded the criterion, the test was performed again at one quarter the contrast; if the response was below the criterion, the contrast was multiplied by 4. The response to the second test was also compared to the criterion and a third test performed at either half or double the contrast of the second test. Thus, with only three tests the response contrast sensitivity could be interpolated to within a factor 2 over a range of 64. These sensitivities were calculated afterwards from the data stored on tape.

The most useful new analytical tool was Fourier analysis of the averaged responses into harmonics of the modulation frequency. This was done by having the computer multiply the averaged response (in impulses/sec) by sine and cosine of each harmonic, and scale the product by the proper normalization factor (2/no. of bins). The amplitude of the harmonic was the root mean square of the sine and cosine components, and the phase shift was tan⁻¹ of the ratio sin/cos. We made the machine calculate the amplitudes and phases of the first ten harmonics of each response. For a linear system with sinusoidal input there ought to be significant strength only in the fundamental. In fact, in Y cells there was often a significant second harmonic amplitude. Two graphic examples of the Fourier analysis of two different averaged responses from one Y cell (one with significant harmonic distortion and one without) are given in Fig. 2.
Fig. 2 Fourier analysis of averaged visual neurone response. In the left column of the Figure are two averaged responses or PST histograms, both from a single Y cell. The stimuli were alternating sine gratings at peak and 90° away in spatial phase with 4 Hz sinusoidal temporal modulation. The spatial frequency of the gratings was 0.14 c/deg. The contrast was 0.04 for the upper response and 0.15 for the lower. The right hand column shows the Fourier analysis of these two responses into harmonic components of the modulation frequency. The peak response is predominantly at the fundamental while the response 90° away from peak has significant second harmonic, and the higher harmonics for both are in the noise.

RESULTS

In this investigation we studied forty-five optic tract axons, 21 Y and 24 X in nine cats. Thirty-four fibres, 19 Y and 15 X, were studied for over 5 hr each, while the remainder were only under scrutiny for about an hour. For all these cells, position in the visual field and centre response type (on or off) were determined by plotting the receptive field by hand with an ophthalmoscope on a tangent screen under a low background illumination (10⁻³ cd/m² approximately).

Linear/nonlinear test: theory

All the optic fibres were classified as X and Y by use of a modified ‘null test’ (Enroth-Cugell & Robson, 1966). The interpretation of this kind of experiment is the cornerstone of this and the accompanying paper (Hochstein & Shapley, 1976), so it is important to make clear what the null test is and what it does and does not measure.

In our experiments, the ‘null test’ was performed with an alternating
phase sinusoidal grating as a stimulus. Occasionally we used introduction and withdrawal of grating (an on-off grating) with no essential differences in results. It should be noted that the space averaged luminance of the stimulus did not vary with time, but that the pattern of spatial non-uniformities in luminance was time-modulated. To study the spatial properties of the ganglion cells, we controlled the position of the grating in the visual field. Since the grating is a periodic stimulus, it is natural to consider the position of the grating in the visual field as equivalent to the spatial phase of the grating with respect to a fixed point in the visual field, e.g. the mid-point of the receptive field of the visual neurone being studied.

Formally, for sinusoidal temporal modulation the alternating grating stimulus luminance was $L_0 + L_1 \sin (2\pi kx + \phi) \sin (2\pi ft)$ where $L_0$ was the mean luminance, $L_1/L_0$ was the contrast, $k$ was the spatial frequency in c/deg, $x$ was the position on the screen in degrees of visual angle, $f$ was the temporal frequency in c/s, $t$ was time, and $\phi$ was the spatial phase. For square-wave temporal modulation, the stimulus was $L_0 + L_1 \sin (2\pi kx + \phi) \text{SIGN} (\sin 2\pi ft)$ where $\text{SIGN} (a)$ is $+1$ if $a$ is positive and $-1$ if $a$ is negative. Fig. 1 shows a graphic representation of the stimulus wave form.

The force of the null test devised by Enroth-Cugell & Robson (1966) comes from the following considerations. In this kind of experiment spatial and temporal frequency are fixed, and so spatial phase is the only variable stimulus parameter. Under these conditions modulation depth at each point of the stimulus becomes a sinusoidal function of spatial phase. To the extent that response at each point on the retina is a linear function of the stimulus, and to the extent that these local linear responses are simply added up in spatial summation, the response of the ganglion cell will also be proportional to a sinusoidal function of the spatial phase. To look at it in another way, given the conditions of linearity stated in the previous sentence, the ganglion cell will act as a linear spatiotemporal filter. For linear filters sinusoidal inputs yield sinusoidal outputs. Thus, the cell's response is a sinusoidal function of spatial phase, and there will be two 'null positions' at which the response goes to zero, the zero crossings of the sine function. This argument hinges only on the sinusoidal wave form of the spatial stimulus, linearity of local responses, and linearity of spatial summation. It does not depend on symmetry of the cell's sensitivity profile and is therefore completely general for any sensitivity profile. Retinal ganglion cells have even-symmetric sensitivity profiles so it is easy to comprehend the sinusoidal dependence of sensitivity vs. spatial phase for a sine grating in these cells. The peak sensitivities occur when the maximum and minimum luminance of the sine grating (the peak and the trough) are lined up with the peak of the receptive field sensitivity profile, and the nulls occur when zero crossings of the sine grating are
lined up with the peak sensitivity. Other spatial patterns besides sine gratings will yield null positions in visual neurones which perform linear spatial summation. To give a null, a spatial pattern must have spatially distinct increments and decrements which can generate equal but opposite responses locally in the retina, which responses then cancel at the level of spatial summation.

It is useful to make a clear distinction between linearity in the temporal domain and linearity in the spatial domain. The latter is what we mean by linearity of spatial summation. Linearity simply means the property

![Graph showing impulse responses for X and Y cells](image)

Fig. 3. Null tests for X and Y cells for slowly flashing (0.5 Hz) gratings. Alternating phase sine gratings at 0.5 Hz square-wave temporal modulation were used as stimuli. These runs are all at 0.15 contrast, but in general the experiment was done at a number of contrasts. The top row of responses belongs to the X cell; the bottom row is from the Y cell. The X cell was, on-centre, 40° peripheral and the Y cell on-centre, 20° peripheral with respect to area centralis. The spatial frequencies were 0.7 c/deg for the X cell, 0.35 c/deg for the Y cell. The impulse rate calibration is indicated, as is the time scale of all the runs. The two peaks of the X cell and Y cell responses are shown as is the 'null' position 90° away from the peak positions in spatial phase, at which position the X cell really gives almost zero response while the Y cell responds at one alternation, then the other (an 'on-off' response). Responses were taken every 30° in spatial phase routinely and automatically, but only the peaks and one null are shown here.

of superposition: the response, to a compound stimulus which is the sum of two simpler stimuli, is the sum of the responses to each of the simpler stimuli presented individually. Superposition implies proportionality of response with stimulus, and it also implies inversion of the response to an
inverted stimulus. For stimuli restricted in space the only linearity that can be tested is temporal linearity, but for distributed spatial patterns, one can and does test spatial and temporal linearity with experiments like the null test. Nonlinearities after spatial pooling do not affect the null test, nor do they affect the sinusoidal dependence of sensitivity on spatial phase of a sine grating. However, such late nonlinearities should be evident in the dependence of response on contrast. Nonlinear stages which occur after local responses but before spatial pooling are exposed by the null test.

The presence of two null positions in the graph of ganglion cell response vs. spatial phase of a sine grating is not sufficient by itself to prove linearity in the processes of spatial summation. If one only possessed the modulated responses of the cells and only had performed the null tests at one contrast, it might be possible to hypothesize a fortuitous cancellation of nonlinearities yielding the existence of null responses. However, by measuring the responses at a few contrasts, and thereby determining sensitivity rather than simply response as a function of spatial phase, one eliminates such a hypothesis. On the other hand, the absence of any null positions is an indicator of nonlinearity at or before the stage (or stages) at which spatial integration takes place.

*Linear/nonlinear test: experiments*

Even though our stimulus conditions were not identical to those of Enroth-Cugell & Robson (1966), it was possible for us to reproduce their basic findings. Fig. 3 illustrates this point. The X cell in the upper row gave maximal responses at two positions of the alternating grating 180° (a half period) apart in spatial phase, and gave a negligible response when the grating was 90° in spatial phase from these maxima. For the Y cell in the lower row there was no position of the grating which gave a negligible modulated response. At 90° spatial phase from the two peak responses, which would be a null position in an X cell, the Y cell gave transient responses at each phase of pattern alternation, similar in time course to the ‘on-off’ responses originally observed in frog retina by Hartline (1938). We will refer to these doubly excitatory responses as ‘on-off’ responses throughout this paper.

Although this method produced a separation of ganglion cells into distinct classes, it gave rise to an obvious question: by using the null test, is one separating the ganglion cells into groups based on a fundamental property of the cells’ wiring, or rather does some fortuitous combination of stimulus characteristics produce the dichotomy? Questions arose concerning the simple X/Y classification when one tried to answer this question by the straightforward manipulation of the important stimulus parameters: temporal frequency, spatial frequency, and contrast. These
questions can be answered by adopting a new classification scheme. In what follows we will first present the results of the experiments on extending the null test in parameter space and the questions raised by these experiments. Based on the results of these experiments, we will present our own proposed classification scheme which helps to answer the basic question, namely whether there are separate types of cells or whether the cells form a continuum.

![Graph showing spatial phase and impulses/sec for X and Y cells](image)

**Fig. 4.** Null tests for X and Y cells for 4 Hz flashings gratings. These data are from the same two cells as in Fig. 3. Again the top row is from the X cell, bottom from the Y cell. Here 4 Hz square wave temporal modulation was used, but the spatial frequencies, spatial phases, and contrast were the same as in Fig. 3. Only one peak is shown for the Y cell. Again the Y cell responded at double the modulation frequency at a spatial phase of the grating 90° away from the peak. The X cell gave essentially no response at a position of the grating midway between the two positions which gave peak responses.

**Temporal frequency and wave form**

The first variable we changed was temporal frequency. Fig. 4 shows the results of a null test performed with 4 Hz square-wave alternation on the same cells as in Fig. 2. At the higher temporal frequency, the Y cell response at 90° spatial phase away from the peak is again ‘on-off’ in nature. The X cell gave a negligible response at a corresponding position of the grating. The response of the X cell is not entirely zero at the 4 Hz null but this is probably due to a very small drift of the eyes relative to the
spatial pattern. The position of the pattern was adjusted initially by hand to get the null at 0.5 Hz and the computer ran the experiment from then on. Small eye movements of the order of 0.1° could lead to departure from a null response of as much as is shown in the Figure. This is one of several weaknesses of the null test experimental paradigm which we sought to improve upon with the sensitivity measurements to be described below. Similar results were obtained when the temporal wave form of the pattern alternation was sinusoidal rather than square wave. Also, there was nothing special about 4 Hz; most temporal frequencies of alternation produced pretty much the same pattern of responses in X and Y cells, with some exceptions. For example, pattern alternation with a slow sinusoidal temporal wave form allowed Y cells to produce more of a null response at a spatial phase 90° from the peak. The reason for this will be considered elsewhere in a detailed report on the dynamics of X and Y cells.

Spatial frequency

Spatial frequency of the alternating grating was the next stimulus parameter to be examined. Profound changes in the pattern of responses in the null tests were produced by altering the spatial frequency. Illustrations of the main effects are presented in Fig. 5. Complete null tests are shown at two different spatial frequencies for one X cell and one Y cell. The alternating grating was modulated sinusoidally at 4 Hz and the interval between spatial phases was 30°. The averaged responses for the peak positions are placed in the centre of each row and at the left and right end. The responses 90° away from these peaks are, as expected, null positions. The main point is that, for each Y cell, the pattern of response changed, becoming spatial phase-insensitive at spatial frequencies high for the cell. Also, the phase-insensitive Y response at high spatial frequency was always 'on-off' and resembled the response obtained at low spatial frequency when the alternating grating was placed 90° away from the peak spatial phase. Both these results were obtained also by Enroth-Cugell & Robson (1966) who used a rather different temporal wave form and temporal frequency.

Another interesting aspect of the effect of spatial frequency is that some X cells began to lose the null position when low spatial frequencies were used for the null test. A hint of this can be seen in Fig. 5, at the null response for the X cell when spatial frequency was 0.14 c/deg. At higher spatial frequencies the response at the null position was more nearly null for X cells. This finding was also adumbrated by a result of Enroth-Cugell & Robson (1966): the minimum response of an X cell to introduction and withdrawal of an edge pattern was not totally absent and showed a distinct 'on-off' response (cf. their Fig. 12). These results are relevant to
Fig. 5 Null tests for X and Y cells at different spatial frequencies. In this Figure two null tests for each of two retinal ganglion cells are shown. The upper two rows are the results of null tests at 0.14 c/deg and 1.4 c/deg for an X cell, while the lower two rows are results at 0.35 and 0.7 c/deg for a Y cell. Sinusoidal temporal modulation of the alternating phase grating at 4 Hz was used, as indicated in the bottom row of the Figure. Every row has peak responses at the ends and in the middle (spatial phases ± 180° and 0° which are really equivalent when using an alternating grating). One quarter cycle away from the peaks are the expected nulls. The 0.7 c/deg series for the Y cells shows no peak response; all the responses are more or less identical. That is the main point. The X cell was on-centre, 30° from area centralis while the Y cell was on-centre 15° from a.c. Contrast was 55% in all cases.
the question whether the centre and surround mechanisms interact linearly in X cells, a subject raised in the Discussion.

The presence or absence of the 'on-off' response at higher spatial frequencies (> 0.5 c/deg for most cells within 20° of area centralis) gave a very reliable indication of whether a cell was X or Y. However, at low spatial frequencies (< 0.2 c/deg for the most cells within 20° of area centralis) the behaviour of Y cells in the null test looked very X-like and some X cells exhibited Y-like behaviour. One question this raised was whether X and Y cells were two ends of a continuum rather than distinct classes. In order to answer this fundamental question we made use of more rigorous techniques: contrast sensitivity curves and Fourier analysis. These new techniques were then used to establish a more solid classification of cells as X and Y.

Contrast sensitivity

Contrast sensitivity as a function of spatial phase was measured by performing the null test at a number of contrasts, obtaining a response vs. contrast curve at each phase, choosing the contrast which gave a criterion response, and taking the reciprocal of this contrast-to-reach-criterion as the contrast sensitivity. Initially we chose the peak firing rate minus the minimum firing rate as the response to be measured. If the criterion was set low enough, below 50 imp/sec say, the peak of the averaged response was simply proportional to contrast. Under these conditions, contrast sensitivity could be given the absolute units imp/sec ÷ contrast; it was the slope of the straight line which fitted the response vs. contrast curve. Contrast sensitivity could be measured at higher criterion levels outside this linear response range, but then the sensitivity could not be expressed in absolute units. Some sign had to be given to the contrast sensitivity. This was assigned somewhat arbitrarily, according to the following convention.

The sign of the contrast sensitivity, plus or minus, was referred to the temporal phase of the response with respect to the temporal phase of the modulation signal which produced the grating alternation. If the response increased when the modulation signal was positive, the contrast sensitivity was given a positive sign; if it increased during the negative phase of temporal modulation, it was called negative sensitivity. This is clearly arbitrary depending on what type the cell was, on or off-centre, and what the starting spatial phase of the spatial grating was with respect to the receptive field centre. But it is an internally consistent and useful convention, particularly for Y cells which had two sensitivities at some positions (see below).

Fig. 6A shows the contrast sensitivity graph for an X cell at two different spatial frequencies and Fig. 6B the corresponding functions for
a Y cell. Sinusoidal curves which are least squares best fit to the X cell contrast sensitivity data are drawn in Fig. 6A. Remember that in a cell with linear local responses and linear summation, the sensitivity should be
a sinusoidal function of spatial phase. The X cell data do not deviate significantly from the sine curve. The Y cell data could not be approximated by sine curve because of double valued sensitivities: at low to medium spatial frequency only near what should be the null, and at high spatial frequency at all positions. The Y cell sensitivities are double valued at positions for which there are peak responses at each half cycle of alternation. Similar curves for X and Y cells in the cat lateral geniculate nucleus have been obtained previously (Shapley & Hochstein, 1975).

Fourier analysis of response

Fourier analysis of the averaged neural responses before the construction of the contrast sensitivity curves was used to extend this analysis one step further. This technique was especially useful when the alternating grating was temporally modulated with a sine wave. A linear cell should only respond at the modulation frequency, while non-linearities should make themselves evident as harmonic distortion, i.e. the presence in the response of modulations at 2f, 3f, 4f, etc., when the stimulus modulation frequency was f. Consistent with previous results, X cells mainly responded at the modulation frequency, although there was some harmonic distortion in large responses produced by very strong stimulation (> 0.1 contrast at the peak of sensitivity). The Y cells showed much more distortion, especially second harmonic distortion, even for low contrast stimuli and small responses. Typical results for an X-Y pair are shown in Fig. 7. The dependence of contrast sensitivity on spatial phase for the fundamental component of the Y cell resembles the sensitivity vs. spatial phase function in the X cell. Second harmonic distortion sensitivity in Y cells was typically of the order of 100 imp/sec/contrast (for low contrasts), and was roughly

Fig. 6. Contrast sensitivity of peak response to alternating grating as a function of spatial phase. The slope of the straight line which related peak-trough of the averaged response to contrast (or the reciprocal of contrast to give a constant response) was used to quantify the dependence of sensitivity on spatial phase. This quantity which we called contrast sensitivity has units impulses/sec ÷ contrast. In Fig. 6A the contrast sensitivity vs. spatial phase is shown for an X cell at two different spatial frequencies (0.14, □ and 0.7 c/deg, ○). The temporal stimulus was 4 Hz square wave. The continuous curves are sine functions which best fit the data by the method of least squares. In Fig. 6B the same function for a Y cell are shown at two spatial frequencies (0.35, × and 0.7 c/deg, ○) and also for 4 Hz square-wave alternation. Note at 0.35 c/deg there is a dependence of contrast sensitivity on spatial phase but it has a double valued peak where it should go to zero. At 0.7 c/deg there is no structure at all; the peak response is phase insensitive and double valued everywhere. No continuous curves were drawn to fit these data. The X cell was on-centre 40° peripheral, the Y cell was off-centre 15° peripheral.
independent of the spatial phase of the stimulus. At high spatial frequencies the second harmonic distortion dominated the response of the Y cell. This was not because distortion grew with increasing spatial frequency but rather because sensitivity to the fundamental frequency diminished. This fact was not obvious when looking at the raw PST histogram or at the contrast sensitivity vs. spatial phase plot for peak responses, but stood out.

Fig 7. For legend see facing page.
clearly through the use of Fourier analysis. Much is made of this finding in the accompanying paper (Hochstein & Shapley, 1976).

The findings on second harmonic distortion in Y cells imply that there is an essential nonlinearity after local responses and before spatial pooling in these cells. A nonlinearity is an essential nonlinearity if it does not go away when the stimulus strength is made smaller. The second harmonic component in Y retinal ganglion cells is bigger than the fundamental component under some stimulus conditions, even at very low contrast. The contrast sensitivity measure also shows that the sensitivity of the second harmonic exceeds the fundamental sensitivity at high spatial frequencies in Y cells, and so must be due to an essential nonlinearity at or prior to spatial integration.

**Classification by nonlinearity index**

The most rigorous classification scheme we could invent is based on the ratio of the second harmonic sensitivity to the peak sensitivity for the fundamental component. Fig. 8 shows how the 2nd harmonic/1st harmonic ratio depends on spatial frequency for one X and one Y cell. At very high spatial frequencies both harmonic components sink into the noise and so the asymptotic value of the ratio is 1.

In classifying the cells, we took the maximal value of the ratio over a range of spatial frequencies and assigned the cell that number, termed it the **nonlinearity index**. A histogram of number of cells vs. nonlinearity index is shown in Fig. 9. The X cells are all in the peak below 1; the Y cells are strung out along the abscissa with values of the index extending from 2 up to 17. Using this criterion for classifying the cells we found no overlap. Furthermore, it was rigorously based on a property of the cells invariant with temporal wave form, temporal frequency, and contrast of the stimulus.

---

Fig. 7. Contrast sensitivity of Fourier components to alternating grating as a function of spatial phase. Again contrast sensitivity was determined but with harmonic amplitude as the criterion. In Fig. 7A, the results from an X cell are plotted. The spatial frequency was 0.7 c/deg. Sensitivity for the fundamental response is marked by ×. Also, the second harmonic (D) is graphed and is seen to be negligible. In Fig. 7B the results for a Y cell are shown. The fundamental component at spatial frequency 0.14 c/deg is signified by ×, at 0.7 c/deg by □. The second harmonic component at 0.14 c/deg is indicated by ·, at 0.7 c/deg by ○. The spatial phase insensitivity and relative spatial frequency insensitivity of the second harmonic component of Y cell responses are the most significant aspects of these graphs. For both cells the temporal modulation was 4 Hz sinusoidal alternation. Best fitting sine curves are drawn for each fundamental component as a function of spatial phase. The X and Y cells were both on-centre; the X cell was 15° peripheral and the Y cell 25° peripheral from the estimated centre of vision.
(As shown above, the original null test does not depend strongly on temporal frequency or temporal wave form. Our refined null test, which involves measuring contrast sensitivity vs. spatial phase, does not depend on contrast.) The nonlinearity index classification scheme was also invariant with retinal locus, as can be seen in Fig. 10. There the positions of the ganglion cell receptive fields in the visual field are plotted, along with the X/Y identification based on the nonlinearity index. X and Y cells are both found near the region of central vision and in the far periphery, consistent with the available histology (Wässle, Levick & Cleland, 1975; Boycott & Wässle 1974).

Some sceptics may persist in believing that peripheral X cells and central Y cells have similar spatial summation properties. This is just not true. The X cell which yielded the contrast sensitivity graph in Fig. 6A was a very peripheral cell while the Y cell in Fig. 5 was much closer to area centralis. Yet the nonlinearity indices were 0.5 for the X cell, 10 for the Y cell.
Fig. 9. Nonlinearity index histogram. Here are plotted all the cells for which we had enough data to draw a graph of nonlinearity index vs. spatial frequency. The nonlinearity index was the maximum value of the 2nd/1st ration. All the X cells are in the peak below 1; all the Y cells had values for this index not less than 2. (One Y cell with a nonlinearity index of 2·0 was plotted in the bin between 1·5 and 2.)

Fig. 10. X and Y cells vs. eccentricity. Each cell is plotted in the visual field with the area centralis of each eye superimposed on the other as in normal fixation. O is the fixation point; L is the left optic disk; R, the right optic disk. The different cell types are indicated by different symbols. The concentric circles with O as centre have radii of 10°, 20°, and 30° in the visual field.
Comparison with other ganglion cell classifications

Having attempted to establish a rigorous linear/nonlinear classification, we also wanted to compare the classification of cells along this dimension with other classification schemes. So on many of the same cells used for the linear/nonlinear tests we used the following tests, originally used by the cited authors: (1) elevation of discharge frequency in response to a drifting grating (Enroth-Cugell & Robson, 1966; Cleland et al. 1971; Hoffmann et al. 1972), (2) spatial frequency tuning (Maffei & Fiorentini, 1973), (3) standing contrast (Cleland et al. 1971), (4) response to diffuse light (Fukada, 1971).

The stimuli we used to perform these tests are briefly described in the Methods section. All stimuli were produced on the same oscilloscope screen as was used for the alternating phase gratings in the null tests, and the mean luminance was kept at 1 cd/m². This luminance is lower than that used by many others, but subsequent experiments in this laboratory using a 20 cd/m² background have produced results consistent with the observations at 1 cd/m² (E. Kaplan, Y. So & R. M. Shapley, unpublished results). The drifting gratings used in test nos. 1 and 2 were produced by an electronic circuit (described in Shapley & Rossetto, 1976) which allowed independent control of spatial frequency and temporal rate of drift (in Hz, or periods of the grating/sec). For measurement of spatial tuning, temporal drift rate was held constant and spatial frequency varied. Contrast sensitivity as a function of spatial frequency was calculated from the responses over a range of contrasts.

For tests nos. 3 and 4, the full screen of the oscilloscope (10° × 13°), 1° to 2° wide bars, and 1° × 1° spots were used as targets. These targets were modulated in a square-wave manner around the mean luminance; at one phase of the square wave the target was brighter than the background and at the other phase it was darker than the background. The contrast, or depth of modulation, was between 0·075 and 0·3. For the standing contrast test, test no. 3, sinusoidal gratings were also used as targets which could stimulate mainly the central response mechanism of the ganglion cells.

Drifting gratings. There was very good correlation of the nonlinearity index with test (1) and (2). That is, an X cell with a low value of the nonlinearity index would not show an elevation in average firing rate when presented with a drifting grating at a high spatial frequency, and the same X cell would have a low spatial frequency attenuation. A Y cell would show the elevation and would not have a low spatial frequency cut-off, down to 0·14 c/deg, as low as we measured. However, the presence or absence of a low spatial frequency attenuation in Y cells was critically dependent on drift rate. At the drift rate used by Maffei & Fiorentini (1973) (4 Hz) we
also found that the Y cells showed no low spatial frequency attenuation, but at slower drift rates a low spatial frequency attenuation often became apparent. In X cells the low spatial frequency roll-off was not so dependent on drift rate. Thus, in order to classify cells by means of spatial tuning to drifting gratings, one needs to control the temporal parameter.

*Standing contrast.* The averaged responses of X and Y retinal ganglion cells to diffuse light and various standing contrasts designed to isolate the central response mechanism are shown in Fig. 11 and 12. The stimulus modulation rate was 0.5 Hz in Fig. 11 and 0.1 Hz in Fig. 12. The data in both Figures are from the same four cells: one paracentral off-centre X cell, one far peripheral on-centre X cell, one peripheral on-centre Y cell, and one peripheral off-centre Y cell. The contrast was 0.3 for all the runs shown in these Figures. The identification of the cells as X or Y was done with the nonlinearity index. This categorization is supported by the peak and null responses of these cells as shown in the four columns to the right of Fig. 11. The Y cells give an ‘on-off’ or frequency doubled response at the null position for the fundamental response. The two X cells give pretty good null responses both at low (0.15 c/deg) and high (0.77) spatial frequency of the stimulus grating. The responses of these cells to standing contrasts and diffuse light were representative of the rest of our sample.

It can be seen from the left hand column of responses in Figs. 11 and 12 that the time course of response to a centrally located modulated spot was not identical in all X cells. The peripheral X cell produced a brief, transient response to the square-wave modulation while the central X cell gave a much more prolonged response. One may notice that the spot response of the (far) peripheral on-centre X cell was more transient than the corresponding response of the peripheral off-centre Y cell.

Spots are not the only kind of standing contrast one can use to try to isolate the central response mechanism. If one chooses the spatial frequencies and position of the grating correctly, a grating can be equally effective for this task. In Fig. 11, the peak responses for the high spatial frequency grating in the cells appear to be generated by the same mechanism as underlies the spot responses, as judged by the similarity in wave forms of the responses. A similar statement applies to the peak response to the low spatial frequency grating in the on-centre Y cell in Figs. 11 and 12.

The main conclusion from our experiments on time course of averaged responses to standing contrasts is that X cells generally produce a more prolonged response than Y cells, but that the response time course is not by itself a reliable index for classification. The response of the receptive field centre of some Y cells had a substantial steady-state component, while the central response of some X cells only possessed a transient component.
Fig. 11. Responses of X and Y cells to flashing spots and diffuse light and alternating gratings at temporal frequency 0.5 Hz. Each row of averaged responses is for a different cell. The top row is for an off-centre X cell which was 6° from the centre of area centralis; the second row is for an on-centre X cell 50° from a.c.; the third row is for an off-centre Y cell 15° from a.c.; the fourth row is for an on-centre Y cell 15° from a.c. with a high maintained discharge. The spot was in all cases a 1° × 1° square. The diffuse light stimulus was really the full screen – a 10° × 13° rectangle. The low SF grating was 0.15 c/deg; the high SF grating was 0.77 c/deg. Temporal modulation was in all cases shown here square wave, as indicated in the Figure. We also performed runs with temporal modulation going between plus and zero instead of plus-minus, with no qualitative differences.
Diffuse light. The response wave form of retinal ganglion cells to diffuse light has also been used as an index of ganglion cell type (Fukada, 1971). Fukada found that some ganglion cells gave ‘on-off’ responses to diffuse light, while others responded only at light on or light off. The former group were called Type I and had several other properties which made them resemble the transient class of Cleland et al. (1971). The cells which gave

![Graph showing responses of X and Y cells to flashing spots and diffuse light and alternating gratings at 0.1 Hz. These runs at a slower temporal modulation frequency are for the same cells and the same spatial patterns as in Fig. 11.](image-url)
only ‘on’ or only ‘off’ responses were called Type II and resembled the sustained class of Cleland. The diffuse light response of the four representative cells are shown in the second columns of Figs. 11 and 12. The central X cell was the only cell which did not give an ‘on-off’ response to diffuse light at 0.5 Hz (Fig. 11). At 0.1 Hz (Fig. 12) again the peripheral X cell and the two Y cells gave responses on and off of illumination, but the on-transient in the on-centre Y cell (third row) was to some extent obscured by the very large surround ‘off’ response. This masking was not only apparent in the averaged response. One could hear it in the impulse discharge on the audio monitor during the experiment. Our conclusion from these results is that classification of ganglion cells by the presence or absence of ‘on-off’ responses to diffuse light is also unreliable, in that X cells may be grouped with Fukada’s Type I class, while some Y cells might be grouped with the type II cells.

DISCUSSION

Categories

One can classify cat retinal ganglion cells into two non-overlapping groups based on the linearity or nonlinearity of spatial summation. This is done by categorizing each cell by an index of nonlinearity in spatial summation which index reaches a maximum value for patterns near the limit of spatial resolution of the cell. There is solid evidence that this modified X/Y classification is independent of receptive field type (on or off) or retinal eccentricity of the receptive field. The distinction between visual cells based on linearity of spatial summation also holds in the cat lateral geniculate nucleus (Shapley & Hochstein, 1975). We feel strongly that a classification scheme should not be based on properties which are graded over the retina and for which all cells take their place in a continuum. There is a gap in the value of the nonlinearity index between the X cells and all the Y cells, so they do not form a continuum with respect to this index.

Other tests for classifying cells correlate well with the nonlinearity index, e.g. the elevation of mean rate in responses to drifting gratings. We have not done the experiment, but suspect that the ‘spinning windmill’ test (Cleland et al. 1973) would also correlate well with the nonlinearity index. However, the time courses of averaged response wave forms to diffuse light, and to standing contrasts which evoke mainly central responses, are not always associated with the X/Y property under the conditions we used. Specifically, peripheral X cells could perhaps be grouped with central Y cells when the diffuse light and standing contrast test are used, and some central Y cells may be classified with the X cells. If a
battery of tests are used, and not all of the tests are invariant with eccentricity, the possibility of confusion exists. Also, it is somewhat misleading to label the cells X and Y when no precise test of the linearity of spatial summation has been performed, e.g. Hoffmann et al. 1972. Underlying this discussion is the assumption that one test which separates the ganglion cells clearly is more physiologically significant than a whole battery of tests none of which disproves the hypothesis of a continuum.

Adaptation

One experimental variable we have not varied systematically is the mean luminance, and thereby the adaptation level. Some preliminary experiments indicate that Y cells retain high nonlinearity indices when the background is raised into the photopic range from our usual 1 cd/m² background. It is already known that the time course of averaged responses is dependent on adaptation level (Yoon, 1972; Enroth-Cugell & Shapley, 1973a, b). A study by Jakiela, Shapley & Enroth-Cugell (1976) showed that X and Y cells had similar time courses of response when they were completely light adapted in the scotopic range, but that Y cells became light adapted at a lower background and also had more transient square wave responses when partially light adapted. The results of Jakiela et al. and the earlier work on adaptation and dynamics imply that our findings of X cells and Y cells with either sustained or transient square wave responses may be partially due to the fact that we kept mean luminance rather than adaptation level fixed. However, not enough is known yet about adaptation and dynamics, and its relation to the X/Y property, to reach a firm conclusion.

Nonlinearity in X cells

It would be a mistake to call the X cells linear cells and the Y cells nonlinear cells. Rather the X/Y classification deals only with linearity of spatial summation. Each cell type may contain linear and essentially nonlinear components. The accompanying paper (Hochstein & Shapley, 1976) analyses the components of the Y cell receptive field. Here we would like to make some brief remarks about nonlinearity in X cells. The centre of an X cell is clearly linear, because the centre is the only component of the field responding at high spatial frequencies and X cells behave linearly in that region, i.e. they have a very clear null position, contrast sensitivity sinusoidal with spatial phase, a linear contrast vs. response range, inverse responses at on and off of stimulus, and little harmonic distortion. However, at low spatial frequencies nulls are harder to get in X cells, particularly peripheral X cells. Also, diffuse light responses can be 'on-off' in peripheral X cells, and this is also an indicator of some kind of nonlinearity. We think the most likely candidate for the nonlinearity is the
surround response mechanism of X cells. The most plausible explanation, given what is known now, is that there is linear spatial pooling within the centre and surround of X cells, but that, in some X cells, there is a non-linearity in the surround pathway after the stage of spatial pooling. Therefore stimuli which activate both centre and surround (like diffuse light) give evidence of nonlinearity in these cells while high spatial frequency gratings, which only evoke a response from the centre, appear to stimulate only a linear mechanism. If the nulls for the centre and surround mechanisms did not occur at the identical position, this late nonlinearity in the surround would show up in a null test experiment with a low spatial frequency grating. See for instance the top row of Fig. 5. Many X cells have good nulls at all spatial frequencies, and reveal little second harmonic distortion in response to diffuse light (cf. top row of Fig. 11). In such cells the surround must be as linear as the receptive field centre.

Retinal wiring

Finally, we would like to offer some conjectures about retinal wiring based on our efforts at using quantitative methods to categorize cells. The centre mechanism of X cells (and probably the centre of Y cells too, cf. Hochstein & Shapley, 1976) fulfils many criteria for linearity: response proportional to contrast (over some range), null positions for alternating gratings, response inversion at on and off of a spatial pattern. This behaviour implies that the local responses summed in the X cell centre must also be linear. In particular, the transductions from receptor responses to bipolar responses to X ganglion cell responses must be linear, spatially and temporally. This may be a clue to the significance of the ribbon synapses from receptors to bipolars and from bipolars to amacrine and ganglion cells. There are indications that conventional synapses may rectify strongly (Katz & Miledi, 1967). In order to get graded information linearly all the way from the receptors to the ganglion cells some mechanism different from the conventional synapse might be required, and it seems reasonable to speculate that the synaptic ribbon seen in electron micrographs would be one of the anatomical substrates of this different mechanism.

This work was supported by grants from the U.S. National Eye Institute, EY 1472 and EY 188. S.H. was supported by post-doctoral fellowship from the U.S. National Institutes of Health and Fight for Sight, Inc., New York City. Also we wish to acknowledge the gift of diallyl-bis-nortoxiferine from Dr William Scott at Hoffman-La Roche, Inc.
REFERENCES


