1972

An Analysis of Spatial Summation in the Receptive Fields of Goldfish Retinal Ganglion Cells

Michael William Levine

AN ANALYSIS OF SPATIAL SUMMATION IN THE RECEPTIVE FIELDS OF GOLDFISH RETINAL GANGLION CELLS

A thesis submitted to the Faculty of The Rockefeller University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

by

Michael W. Levine, B.S., M.S.

Approved for publication:

J. Abramson
Assistant Professor

March 15, 1972

The Rockefeller University
New York

PREFACE

In writing this thesis I have received so much encouragement and support from so many people that it is difficult to know how to thank all of them, or where to begin. I am grateful to Dr. Detlev Bronk, President Emeritus of The Rockefeller University, who initiated the unique and stimulating graduate program of which I was a part; also to Drs. H. K. Hartline and Floyd Ratliff, in whose laboratory this work was done. I thank all the members of the laboratory for useful discussions and helpful criticisms; I particularly wish to mention my fellow students John Tuttle and Bob Schor, as well as Bruce Knight, whose assistance with the mathematical parts of this work was indispensable. I thank Louis Nagy, who fed the fish, and Maria Lipski, who cut the red tape. I am also indebted to the many people in the various shops and services of the University, particularly Ray Martin of the Computer Laboratory and Mike Rossetto and Norm Milkman who helped with electronics.

The figures in this thesis were prepared by Ruth Mandlebaum and her assistants in Graphic Services. Much of the typing of the drafts was done by Marie Azzarello; this final copy was typed by Mary Costello. I am also grateful to my wife, Jane, for considerable typing and proofreading (as if that were all she has done for me!).

I am extremely fortunate in that my research adviser was Dr. Israel Abramov. Issy and I worked closely throughout this project from designing and building the apparatus, through experimental design and data collection, to data analysis. It is impossible to catalog how much influence he has had on me; perhaps an indication is that I too am now a cigar smoker.

Support for this work came from the following grants: Research Grant EY 188 from the National Eye Institute, U.S. Public Health Service; Research Grant GB 6540 from the National Science Foundation; and Training Grant GM 01789 from the National Institute of General Medicine, U.S. Public Health Service.

ABSTRACT

The axons of retinal ganglion cells convey the signals from which the brain constructs its visual percept of the world. The considerable processing done in the retina is implicit in these signals; it is therefore desirable to understand this processing in some detail.

Only one aspect of retinal processing is considered--the spatial interactions within either the center or the surround of the receptive field of the ganglion cell. A generalized model for this restricted processing is proposed: the independent signals in each of two separately stimulated areas may undergo non-linear transformation; the two areas may interact at one or several levels, ultimately converging on a final common pathway (the ganglion cell); a further non-linear transformation may be performed on the combined signal in the final common pathway.

To examine the features of such a model, two separate areas within the center of a receptive field are stimulated, both individually and simultaneously. A wide range of intensities is used. Two methods of analysis are applied to these data; each is intended to elucidate particular features of the model.

The first method of analysis, the method of response-summation, compares responses when each spot is illuminated separately with responses when both spots are presented together. A graph is presented of the responses to simultaneous stimulation (physiological sum) as a function of the arithmetically summed responses to each spot alone at the same intensity; the independent parameter is intensity. The plot (response-summation) is not affected by non-linearities which occur before the first interaction of the two areas; if there are no non-linearities at or after the first interaction this plot must be a straight line of unity slope. If there are non-linearities at or after the first interaction of the areas, some other function will be obtained; an analysis is given of the expected forms of this curve for a number of possible non-linearities which might reasonably be postulated.

The second form of analysis, the method of sensitivity-summation, considers the intensities required to elicit a particular response in each area stimulated alone as compared with the intensity required for the same response when both are stimulated simultaneously. A graph is presented (sensitivity-summation plot) showing the logarithm of the summed individual sensitivities as a function of the logarithm of the sensitivity to both spots; the criterion response level is the independent parameter. This plot is not affected by any non-linearities after the signals are combined; if there are no non-linearities before the final combination of responses from the two areas, the resultant plot must be a straight line of unity slope passing through the origin. If there are non-linearities at or before the final combination some other function will be obtained; an analysis is presented of the expected forms of this plot for a number of plausible non-linearities.

The response-summation plot is affected only by non-linearities at or after the first interaction of the two areas, while the sensitivity-summation plot is affected only by non-linearities at all levels up to that at which the signals are finally combined; thus the concurrent application of both methods to the same data effectively divides the processing into a three stage model. Non-linearities before the areas interact affect only the sensitivity-summation plot, non-linearities in the final common pathway after the final combination affect only the response-summation plot, while non-linearities in the region from the first to the final interaction must affect both plots.

The methods of response- and sensitivity-summation are applied to data from single ganglion cells in the excised retinae of goldfish. Chromatic interactions are eliminated by the use of near-infrared lights which stimulate only the long-wavelength receptive cones. Spatial interactions within either center or surround of the receptive fields of ganglion cells are examined. The following minimal model is derived. The receptors in each area perform a non-linear transformation; this may best be described by a square root function, but one which fails at the highest and the lowest intensities. There is also a threshold in

each pathway which must be exceeded before the signal from that area will be included in the ganglion cell response. The signals from separate areas combine at a single level by simple linear summation. The final common pathway includes a non-linearity which may best be described as "compressive"--linear for low levels of response, but gradually reaching a limiting value at high levels. No lateral interactions need be postulated; various attempts to demonstrate a level of lateral interactions other than simple spatial summation all fail.

The time courses of the responses are also examined by applying the same summation methods to short time segments of the responses. A supplementary method (relative timing) is also applied; in this method the responses to a brief test flash are measured when the test flash is presented at various times relative to a longer conditioning stimulus in another area. These analyses lead to the hypothesis that the decay in firing rate as a function of time after the onset of a stimulus is a part of the same process which gives the compressive function after summation.

TABLE OF CONTENTS

Page

LIST OF FIGURES

INTRODUCTION

Genesis

In the early 17th century, the astronomer Kepler studied the optics of the eye and concluded that the light sensitive part of the eye must be the retina. Since that time, there has been a mounting attempt to understand the operation of the mysterious layer of tissue primarily responsible for the sense of sight.

In the ensuing three centuries since Kepler, most of the advances in vision research were made by physicists defining the properties of the light with which the visual system operates, psychologists defining general properties of the system, and anatomists defining the physical appearance of the system. Work on the functioning of the retina had to await the development of electronic instrumentation. In 1927, Adrian and Matthews recorded the "all-or-nothing" action potentials which indicate activity in the optic nerve of the eel. They hoped to single out the message from a single fiber by optically isolating its input with an extremely small stimulating spot, but found that no matter how minute a spot they used, many cells responded; moreover, increasing the spot size led not only to an increase in total response, but a decrease in latency. Clearly, the retina was doing something far more complicated, and more interesting, than simply transmitting a point by point replication of the stimulating light pattern.

The first recordings of single fibers in a vertebrate retina were made by Hartline in 1938. He teased apart the bundles of fibers coursing over the vitreal surface of a frog retina, until his electrode sensed a single active axon. Probing with small spots of light, he found a large area in which a cell could be stimulated, which he termed the cell's receptive field. He also found that different cells had quite different response characteristics; a few continued to respond throughout the duration of a maintained light, but others gave bursts of response only to an increase in illumination ("on" type), others responded only to the reduction of light ("off" type), and still others to either increase or

diminution of light ("on-off" type).

The concept of a receptive field was further extended by Kuffler in 1953. Using a fine platinum-iridium wire with glass insulation at all but the very tip, he recorded from single ganglion cells in the retina of the cat. He found, like Hartline, that stimulation over a roughly circular area (which he now could demonstrate was approximately concentric with the ganglion cell body) evoked some typical discharge, either "on" or "off." However, each cell had an additional concentric region in which the evoked responses were of the opposite type; thus, if the central region of a cell responded with "on" firing, the surrounding area would respond at "off," while intermediate regions would give mixed responses ("on-off"). The outermost zone has become known as the antagonistic surround.

The receptive fields of ganglion cells testify that the retina performs a considerable amount of spatial and temporal information processing, but there is another form of information abstracted by the retina. Many animals have the ability to distinguish wavelengths of light, independently of the relative intensities at the different wavelengths. To do this requires at least two independent mechanisms of different spectral sensitivities[g]; animals with a well-developed color sense, such as man, have three different cone photopigments as the basis of trichromatic[g] color vision. But is there a point by point representation of wavelength, each ganglion cell having one of the available spectral sensitivities? The answer is no, and as in the spatial case, there is an antagonism of mechanisms (Svaetichin and MacNichol, 1958). De Valois and his colleagues (1955, 1966) investigated the spectral responses of lateral geniculate units in the macaque monkey and found them to respond "on" to lights at one end of the visible spectrum, and "off" at the other end of the spectrum. Such

---

[g]The superscripted g shall indicate terms which are defined in the glossary, Appendix I.

responses may be described as "spectrally opponent." Furthermore, the long wavelength sensitive and short wavelength sensitive mechanisms contributing to the response are typically in separate parts of the concentric receptive field of the ganglion cell (see for example Wagner, MacNichol, and Wolbarsht, 1960, 1963; Motokawa, Yamashita, and Ogawa, 1960; Wiesel and Hubel, 1966; Daw, 1967).

The above description constitutes an outline of the generalized receptive field of a retinal ganglion cell. If one were merely interested in the form of the information being sent to the brain, a quantified version of this picture would suffice. But it would be much more satisfying to be able to understand how the retina performs this processing. A functional model of the formation of the receptive field would be valuable both as a better description of the receptive field, and as an example of a piece of brain at least partially understood. The goal of this thesis is to develop a minimal model of the processing underlying a ganglion cell receptive field.

The Goldfish Retina

The experimental work reported in this thesis was performed on the isolated retina of the common goldfish, Carassius auratus. There were a number of reasons for selecting this particular system for study, including considerable previous work on the goldfish (and its close relative, the carp), a number of physiological advantages, and some operational considerations. This section describes the particulars of the goldfish retina and some of the advantages for which it was selected for this study.

Anatomy The goldfish is a fresh-water teleost, and therefore evolutionarily quite distant from the mammals. However, there are remarkable similarities between the fish retina and our own. Like man, goldfish has a duplex retina; that is, it has two receptor systems, a rod system for dim light vision, and a cone system for brighter lighting. Like man, the goldfish has demonstrably trichromatic photopic[g] vision (see for example, Yager, 1967), and goldfish cones each contain one of

three possible pigments. The analogy extends to the deeper layers of the retina as well; the same basic cell[g] types and interconnections as are found in man are found in goldfish (De Testa, 1966).

There are, of course, important differences between the two species. Nearly half the cones in goldfish are paired as twin cones, which are not seen in mammals. Twin cones are pairs of cones with closely contiguous inner segments; one is maximally receptive to longer wavelength light, the other to shorter wavelength light (Marks, 1965). According to Walls (1942), the development of the teleost cone and the mammalian cone were two separate events in evolution, so there may be significant differences in the output of the two types of cones, or in the ways in which the deeper retinal layers make use of cone information. Another difference is in the horizontal cells, of which the goldfish has three distinct layers, including an innermost layer consisting of long tubelike processes not seen in man (Stell, 1967). The amacrine cells in the lower vertebrates are more numerous than in man, and appear to synapse extensively with each other (Dubin, 1970); in man there is a relatively higher ratio of direct bipolar to ganglion cell contacts (Dowling, 1970). The processing in the teleost retina may in fact be more complex than in man's retina, including functions that in man have been subsumed by the brain.

Photopigments The cone pigments in the goldfish were studied microspectrophotometrically by Marks (1965). The transmission of single cone outer segments was measured, and compared with the same measurement after all the pigment had been bleached. The density difference spectra[g] thus obtained are a measure of the photopigment absorption as a function of wavelength; Marks found three distinct pigments, only one in each cone, each of which could be fitted by a Dartnall nomogram[g] indicating a lawful visual pigment (Dartnall, 1962; Munz and Schwanzara, 1967; for a more complete discussion, see Abramov, 1972). The peak sensitivities of the three pigments were at about 455 nm, 530 nm, and 625 nm (Marks, 1965; Liebman and Entine, 1964). For brevity, these will be referred to as "short," "medium," and "long" pigments, respectively, despite the fact

that each is capable of some absorption over most of the spectrum. In 1967, Tomita, Kaneko, Murakami, and Pautler succeeded in recording electrophysiologically from single cones in carp, and corroborated Marks' findings.

It is worth noting that the "short" and "medium" pigments in goldfish have absorption peaks essentially identical to two of the human cone pigments (Marks, Dobelle and MacNichol, 1964; Brown and Wald, personal communication) but the third is sensitive at considerably longer wavelengths than the corresponding human "long" pigment, which peaks at about 560 nm. This means that at longer wavelengths, greater than 650 nm, the goldfish "long" pigment is still highly sensitive, while the other two are attenuated to negligible levels. This is not the case in the primate, where virtually any wavelength of stimulating light that has an effect on one pigment will also have some significant effect on another pigment. In the goldfish it is possible to stimulate with a deep red light (near infrared to human eyes) and stimulate the "long" system exclusively, thereby examining a single color mechanism free of any chromatic interactions. This, of course, is an important advantage of the teleost retina as an experimental preparation.

Receptive Fields of Ganglion Cells The receptive fields of color-coded cells in the macaque monkey as studied by Wiesel and Hubel (1966) have the same simple center and surround structure reported by Kuffler (1953) for the cat, with the significant difference that the center and surround have different spectral sensitivities. Thus, one cell might have a "red-on" center and "green-off" surround, another might have a "red-off" center and "green-on" surround, another would have a "green-on" center and "red-off" surround, and so forth. Other units are spectrally opponent but lack any spatial opponency, while still others lack any spectral opponency.

Wagner et al. (1960) examined the receptive fields of goldfish ganglion cells using small spots of colored light. They found a spectrally opponent center, with only a small spectrally non-opponent

surround: "red-off," "green-on" in the center, "green-on" in the surround (Fig. 1A). This work was extended by Daw (1967), who used annular stimuli to explore the surround. A small number of the cells he found lacked an antagonistic surround, exactly as described by Wagner et al. (1960), but the majority of his units were considerably more complicated; these cells were spectrally opponent in both center and surround, as well as spatially antagonistic (Fig. 1B). Thus a unit might have a "red-on" and "green-off" center, and a "red-off," "green-on" surround, or vice-versa. Such units are known as double-opponent cells. It is possible that there may have been a band of green sensitive region between center and surround, so these units might simply be cells of the type described by Wagner et al. with an additional surrounding area. Daw also reported a cell type that was not spectrally opponent, with a spectral sensitivity in both center and surround roughly the same as the sensitivity of the long wavelength pigment. However, by using intense chromatic adaptation, he was able to demonstrate that some of these cells did indeed have an opposing input from a short-wavelength mechanism that apparently had simply been overshadowed by a far more sensitive long-wavelength mechanism.

It thus seems possible that all three cell types found by Daw might in fact be members of the same class of units with particularly mismatched sensitivities of certain of the contributing mechanisms. A spectrally non-opponent unit could result if one chromatic mechanism dominated; a spatial imbalance could lead to units apparently lacking an antagonistic surround. If this is so, there are only two types of unit in the goldfish: "red-on," "green-off" center with "red-off," "green-on" surround; or "red-off," "green-on" center with "red-on," "green-off" surround; for brevity, the latter may be called "red-off" (-R+R) and the former "red-on" (+R-R). (See Spekreijse, Wagner, and Wolbarsht, 1972, for greater detail.)

The size of the receptive field of a goldfish ganglion cell is quite large. Daw (1967) reported that only center type responses could be found within an area with a diameter of about 1 mm; the surround

Fig. 1. Organization of the receptive fields of spectrally opponent goldfish ganglion cells. A: The receptive field as inferred from mapping with small spots of light. The cell receives inputs from an excitatory mechanism which is most sensitive to short wavelength light (green) and from an inhibitory mechanism which is most sensitive to long wavelength light (red). The graph below shows how the sensitivity of each mechanism might vary across the receptive field. A small spot of red light in the center would stimulate the inhibitory mechanism more than the excitatory, eliciting an "off" response; similarly, a small spot of green light anywhere in the field would stimulate the excitatory mechanism more than the inhibitory, thereby eliciting an "on" response. On the top is shown a map of the areas giving rise to "off" and "on" responses to red and green lights, respectively. B: Receptive field of a similar unit, as inferred by mapping both with spots and with annular stimuli. It is possible that a field such as the one shown in A forms the center of the receptive field in B. C: Oscilloscope traces of the responses of a single cell of the type indicated in B. The shape of the stimulus is indicated to the left of each record; wavelength of the stimulus is beneath each record. The duration of the stimulus (1 sec) is indicated by a horizontal bar along the abscissa. From Abramov and Gordon, 1972.

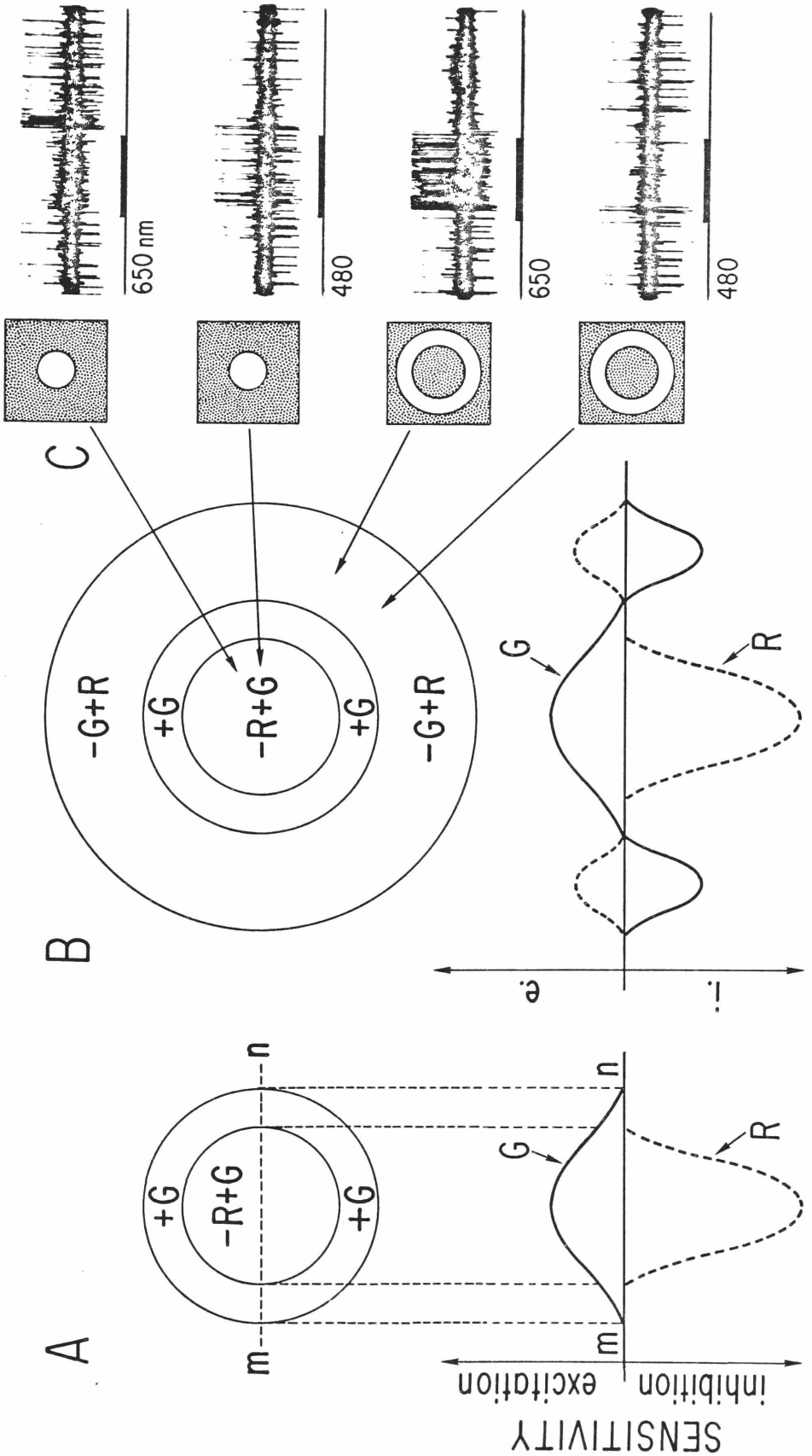
650 nm
480
650
480
C
B
-G+R
+G
-R+G
+G
-G+R
G
R
e.
i.
A
n
m
+G
-R+G
+G
G
R
n
m
inhibition
excitation
SENSITIVITY

Fig. 1

extended to approximately 5 mm, which is about one quarter of the area of the entire retina. Easter (1968a), mapping only the center of "red-on" cells with small spots of red light, found an area of high sensitivity about 0.8 mm in diameter, but there was some sensitivity of the central "red" mechanism within an area nearly 1.5 mm in diameter.

Physiological Considerations There are a number of advantages of working on goldfish which result from the fact that it is cold-blooded. It is possible to remove the retina from the eye and maintain it with a minimum of difficulty. The retina can be kept alive for hours in a slightly chilled state, with only a supply of humid oxygen to nourish it, flush away wastes, and bring enough moisture to prevent drying. Some advantages of the isolated retina will be discussed further in the Materials and Methods section of this thesis.

Although Daw (1967) cited the avoidance of the morass of anesthesiology as an advantage of the isolated retina preparation, recent work by Abramov and Levine (1972) indicates that the problem has not been completely sidestepped. The traditional method of maintaining the retina advocated by MacNichol and Svaetichin (1958) is to introduce a stream of humidified oxygen mixed with 5% CO_2. This prescription was followed by Daw and by Easter, and probably most other workers, though it is often not mentioned explicitly. Abramov and Levine demonstrated that the presence of CO_2 is an important parameter in determining the retinal response. In particular, if the CO_2 is omitted and the retina maintained with a flow of pure humid oxygen, the cells are more sensitive, normally exhibit spontaneous firing in the dark, and, most significantly, seem to be nearly always spectrally non-opponent. In fact, a spectrally non-opponent cell may be converted into an opponent one by simply introducing 2% CO_2 into the oxygen stream. If the retina is maintained with pure O_2, which was the preparation used in these experiments, the non-opponent type seems to be the normal cell. This is a surprising observation in view of the goldfish's demonstrable color vision, but was also found in the intact goldfish by Jacobson (1964) and Adams (1971, and personal communication).

Temperature is also an important parameter affecting the relative spectral sensitivities of the chromatic mechanisms in goldfish. Lower temperatures tend to depress the short-wavelength sensitive mechanisms observed electrophysiologically (Spekreijse, Wagner, and Wolbarsht, 1972); an analogous observation was made by Thorpe (1971), who studied behavior of goldfish stored at various temperatures. The temperature at which the experiments of this thesis were done (13^{o} C) is approximately that used by previous investigators of goldfish retinal physiology. Both the relatively low temperature and lack of CO_2 tend to depress the short-wavelength sensitive mechanisms, and thereby help isolate the long-wavelength sensitive mechanism.

Linearity and the Receptive Field

Considerable work has been done on the organization of receptive fields and the interactions of the underlying mechanisms. I shall consider some of this work here; however, this section is not intended as an exhaustive review but, like the preceding sections, merely a summary intended to convey some of the pertinent facts.

If there were no non-linearities in the retina, the receptive field of a ganglion cell would be completely describable with a single linear equation. As there are non-linearities, an understanding of the receptive field depends on an understanding of these non-linearities, their natures and their interrelationships.

Two points must be borne in mind when considering non-linearities. Firstly, any non-linearity introduced at an early stage in the retinal processing must be expressed in all subsequent stages. Conversely, a non-linearity observed in a particular unit need not have been introduced by that unit, but may have been implicit in the input to the unit.

The second point is that most non-linear functions observed in biological systems may be closely approximated (over a short range) by a linear function. Given a short enough segment of a curve, a secant is an adequate description; before we may conclude that a function is linear we must be cautious that it has been observed over a wide enough

range of inputs to permit any non-linearity to express itself.

Receptors First, let us consider the receptors themselves, as they must be the primary inputs to the receptive field. It is extremely difficult to record from vertebrate receptors, and only a limited amount of information is available. While the spectral sensitivities of the receptors are known from the absorption characteristics of the pigments, it is not clear how the receptors translate captured photons into signals. The receptor is hyperpolarized by light (Tomita et al., 1967; Tomita, 1970), and must have a saturation level which it cannot exceed. According to Werblin (1971), receptors (in Necturus) have a compressive[g] intensity-response function, but do not show sensitivity changes in the presence of steady background illumination. Brown and Watanabe (1965) reported that the receptor potential of the monkey was not greatly affected by long-term exposure to light.

The amplitude of the receptor response in most species depends solely on the intensity of light incident upon that receptor (Tomita, 1970; Werblin and Dowling, 1969). The turtle is exceptional in this respect, for the responses of single turtle cones depend to a large extent on the spatial configuration of the stimulus (Baylor, Fuortes, and O'Bryan, 1971). A specific search for such effects has demonstrated that this is not the case in the goldfish (Tomita, personal communication).

Studies of the input output relationships of the cone can be done by using the massed response of many cones, the late receptor potential (LRP) of the electroretinogram (Brown, Watanabe, and Murakami, 1965). In such a study on the isolated LRP in macaque, Boynton and Whitten (1971) reported that the relationship may be approximated by a (modified) power function, to wit:

$$R = \frac{v \cdot I^n}{k + I^n} \qquad (1)$$

where R is response, I is light intensity, v and k are constants related to the level at which response saturates, and the exponent n has a value which they found to be 0.73. Werblin's results from Necturus and Tomita's

from goldfish are in accord with this equation. However, the data may be approximated by a variety of other functions; if the modified power function is plotted on the same coordinates as either a simple power function with an 0.5 exponent or a logarithmic transform, it can be seen that the differences among them are relatively small (Fig. 2).

Inner Nuclear Layer Less is known of the units between the receptors and ganglion cells. According to Kaneko (1970) both spectrally non-opponent (L-type) and spectrally opponent (C-type) S-potentials[g] may be recorded from each type of goldfish horizontal cell, though this does not necessarily mean that all reported S-potentials have in fact been intracellular horizontal cell recordings. Nonetheless, it is clear from Kaneko's work that horizontal cells have large receptive fields, and may exhibit spectral opponency. As to the intensity-response functions of horizontal cells, S-units recorded by Naka and Rushton (1966) obey $R = I/(k + I)$. This relationship is difficult to reconcile with the observation that the horizontal cell's response to a stimulus is determined by the product of stimulus area and intensity (Norton, Spekreijse, Wolbarsht, and Wagner, 1968), unless one assumes differences in sensitivity across the receptive field of the horizontal cell (see discussion of Ricco's Law under "Ganglion Cells: Interactions within the Center," below).

Bipolar cells are the first level at which a center-surround organization is evident (Kaneko and Hashimoto, 1968; Werblin and Dowling, 1969; Kaneko, 1970; Werblin, 1971). The surrounds of the receptive fields of bipolar cells are of greater area than the spread of the bipolar dendritic arborization (Kaneko and Hashimoto, 1968) implying inputs from horizontal cells. Werblin has also indicated that the bipolar's adaptive state is controlled both by overall luminance and the luminance difference between the center and the surround of its receptive field.

Fig. 2. Comparison of three non-linear functions which have been suggested to describe the responses of the receptors. The ordinate is output magnitude (linear scale); the abscissa is input magnitude (logarithmic scale).

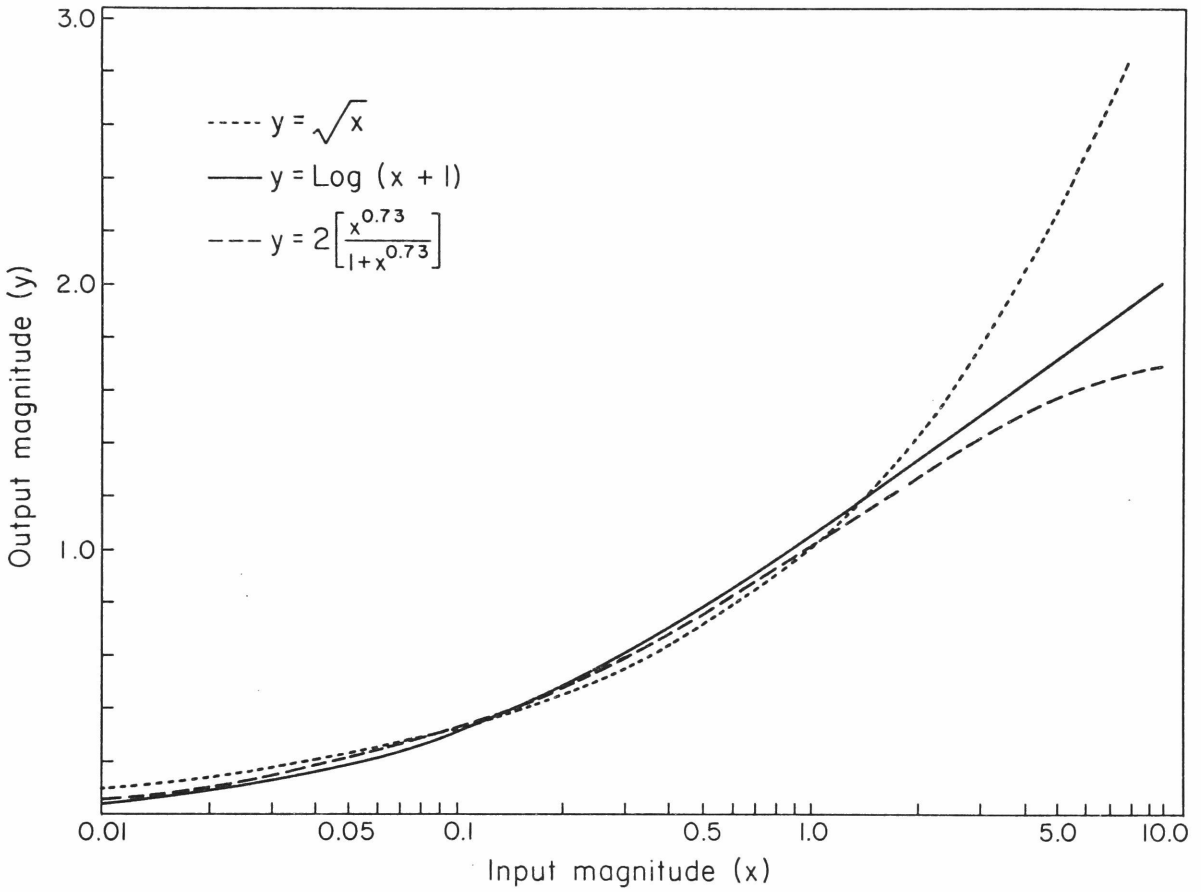
3.0
2.0
1.0
Output magnitude (y)
----- $y = \sqrt{x}$
—— y = Log (x + 1)
--- $y = 2\left[\frac{x^{0.73}}{1+x^{0.73}}\right]$
0.01
0.05
0.1
0.5
1.0
5.0
10.0
Input magnitude (x)

Fig. 2

Amacrine cells are the least understood of all. Werblin and Dowling (1969) found a slight center-surround antagonism in Necturus, as one would expect from the bipolar responses, but Kaneko found no such distinct organization in goldfish amacrines. The intensity-response characteristics of amacrine cells appear steeper than those for bipolar cells (Werblin, 1971). Lastly, it appears that amacrine cells respond quite phasically to changes in the stimulus intensity (Werblin and Dowling, 1969; Kaneko, 1970). This is in contrast to the more distal units, which respond with a sustained potential throughout the stimulus (Kaneko, 1970, 1971; Tomita, 1970; Werblin and Dowling, 1969; Werblin, 1972).

Ganglion Cells: Combination of Center and Surround The ganglion cell receptive field is not only the easiest to study, but perhaps the most interesting. The more distal units are interconnected in complicated ways, and it is not easy to decipher the effect a given cell might have on the next unit down the line; for that matter, it is not yet possible to state which unit is the predecessor of some other. But the ganglion cell presents us with a bottleneck: all information going to the brain must be transmitted by the ganglion cells. They stand in a unique position as the end result of all retinal processing. Whatever combinations and interrelations between groups of receptors and other cells occur in the retina, to be effective they must have happened by the time the ganglion cell spike train is generated.

A considerable amount of work has been done on the summation of the center and surround in the receptive field of ganglion cells in the cat. To a first approximation, the center and surround sum linearly; the response to a spot in the center and the response to an annulus in the surround, if summed point by point in time, match the physiological response to the spot and annulus presented simultaneously (Enroth-Cugell and Pinto, 1970, 1972a, b).

Maffei (1968) used a somewhat different method to demonstrate nonlinear adaptational interactions between center and surround in the cat.

He placed a spot of light in the center of a receptive field; the intensity of the spot could be modulated[g] about its mean level according to a 1/2 Hz sine wave, and the amplitude of the sinusoidal response of the ganglion cell measured. (In assuming a sinusoidal input will yield a sinusoidal output, one is tacitly assuming a linear system, and effectively using linear systems analysis, as discussed in Appendix IV.) A second spot of the same size and mean luminance was held at fixed intensity, but could be moved to different locations in the receptive field. If the system were linear, the amplitude of the sinusoidal portion of the cell's response would be unaffected by the unmodulated spot. Maffei found the amplitude of response depressed when the second spot was in the center of the receptive field, and enhanced when it was moved into the surround. In effect, the second spot changed the state of adaptation of the cell, so it was operating on a different portion of its intensity-response curve. The amplitude of the modulation was so low (10%, which is less than 0.1 log unit) that the curve could be treated as a linear segment, the change in effective mean level yielding a difference in slope (gain). That the second spot can lower the gain when it is in the center is not surprising; that it need not be physically superimposed on the flickering spot implies an adaptation pool, rather than individual receptor adaptation (see Rushton and Westheimer, 1962; Rushton, 1959). It is more surprising that the steady spot in the surround can enhance the sinusoidal response: this could imply a simultaneous contrast mechanism of the adaptive pool, or it might be a result of the overall change of excitation of the unit changing a gain after summation of center and surround.

Maffei, Cervetto and Fiorentini (1970) did the same experiment, measuring the effect on the entire temporal modulation transfer function (MTF)[g]. The general results are in agreement with those of Maffei described above, but depend in detail on the frequency of modulation.

In a very similar experiment on cat, Maffei and Cervetto (1968) found evidence for linear summation of center and surround. They had a spot centered on the receptive field and another spot in the surround;

the intensity of either could be sinusoidally modulated at 1/2 Hz. Treating response as a complex number (sinusoidal amplitude and phase), they showed that the response due to flickering both spots equalled the vector sum of the responses due to flickering each separately, implying linear summation. Note, however, that the range of this modulation was under 0.2 log units.

A similar technique was brought to bear on the problem of the summation of the four underlying spatial and chromatic mechanisms in the goldfish receptive field (see "Receptive Fields of Ganglion Cells," above). Spekreijse (1969), recording from isolated goldfish retinae, found two distinct types of cell with respect to temporal characteristics: the rarer tonic cell, which he did not analyze, responded with a maintained discharge proportional to stimulus intensity; the phasic cell gave a response proportional to the time derivative of the stimulus. Phasic cells fired only to signal changes, having no maintained or spontaneous firing; moreover, they rectified, responding only to increments in the stimulus (or only to decrements, for responses to those mechanisms which yielded "off" responses). In response to sinusoidal stimuli, these cells responded with a clearly half-wave rectified response, with the negative part of the cycle truncated at the zero firing level. It is possible that no spontaneous firing was observed by Spekreijse because he used O_2 with 5% CO_2 (see "Physiological Considerations," above).

Spekreijse wished to use linear analysis, but the presence of a rectifier clearly made the system non-linear. However, by assuming the non-linearity was of a known type, namely a rectifier, and taking its properties into account, the system could be analyzed. First, it was necessary to demonstrate that the remainder of the system was quasi-linear. If the non-linearity were a rectifier with <u>no bias</u>, one could extrapolate the negative side of the response waveform as a mirror of the positive side; the amplitude of this extrapolated, and now sinusoidal, output would be twice the peak to peak amplitude of the actual response. The criterion for linearity is $R_{\alpha x} = \alpha R_x$ (see Appendix IV), where α is a

scalar and R_x is the response to an input of magnitude x. This relationship he found to hold over the limited range of modulation used, up to about 0.7 log units.

Assuming the rectifier to be the only non-linearity, it was then possible to ask whether the rectification occurred before or after various mechanisms were summed. Two stimuli could be focused in the receptive field of the ganglion cell such that each stimulated a different mechanism. By observing the cell's response to the temporal modulation of first one and then the other stimulus, the relative phases of modulation could be adjusted such that the responses would be exactly one half cycle apart; both stimuli could then be modulated simultaneously with this phase relationship. If summation occurred after rectification, the summing would be of two half-wave rectified sines, and the result (if the signals were exactly out of phase at the summation point) would be a full-wave rectified sine. If the summation occurred before rectification, however, and the signals were exactly in counterphase, the two would be cancelled by the summation, and no response would ensue.

The experiment described above was performed by Spekreijse and van den Berg (1971); their finding was that a red spot in the center and a red annulus in the surround, with their phases arranged as described, cancelled nearly perfectly; center and surround summed before rectification. Spekreijse and van den Berg also extended this procedure to examine the summation of the "red" and the "green" mechanisms within the center of the field. Using superimposed temporally modulated lights of 500 nm and 650 nm wavelength, with the phases of the modulations appropriately adjusted, they found a completely null response. Fearing that the interaction might be within the pairs of twin cones, they focused the 650 nm and 500 nm lights on two separate areas within the center; again, there was no response when they were flickered in the appropriate phase. As with center and surround, the "red" and "green" mechanisms sum before rectification.

Ganglion Cells: Interactions Within the Center Just as one might consider how the center and surround of the receptive field of the ganglion cell are summed, one can ask how individual sub-areas within either center or surround combine with each other. Spekreijse and van den Berg found that the "red" inputs from two distinct areas in the center are summed before the rectifier.

The problem of interactions of areas within the center was examined in detail by Easter (1968a), also for the goldfish. Easter first demonstrated a well-known psychophysical relationship, Ricco's law[g]. Working only in the centers of "red-on" goldfish ganglion cells, he found an intensity-response relationship for each of several long wavelength circular stimuli concentric with the receptive field. By interpolating on these curves, he found the intensity (I) needed for each spot to elicit a criterion response (5 impulses); the log of this intensity was plotted against the log of the spot diameter (D). The resulting points fell along a line with a slope of -2, that is:

$$\log I = -2 \log D + k$$

so:

$$I = k'/D^2$$

or:

$$I \cdot A = k''$$

where A is area, and the k's are constants. The last relationship, intensity times area is constant, is Ricco's law.

One interpretation of Ricco's law is that the sensitivity within the area for which the relationship holds is independent of the areal distribution of the stimulus. Thus, only total incident light is significant for excitation, and there is "complete spatial summation." Easter tested this assertion by stimulating with two distinct, well separated spots of light, but both within the area for which Ricco's law held. The two spots were in equisensitive positions, their individual intensity-response curves superimposing quite closely. Easter reasoned that stimulating with both at once should be equivalent to doubling spot area in Ricco's law. Defining I_s as the intensity required

in either individual spot for an arbitrary criterion response, and I_d as the intensity required in each spot for the same response with both illuminated simultaneously, if A is the area of each spot (the area of both is obviously 2A):

$$A \cdot I_s = 2A \cdot I_d$$

$$I_s = 2I_d$$

$$\log I_s = \log I_d + \log 2 = \log I_d + 0.3. \qquad (2)$$

Equation (2) should hold for any criterion chosen. When the experiment was performed, however, a plot (parametric with criterion) of $\log I_s$ as a function of $\log I_d$ showed a straight line with unity slope and an intercept of about 0.6:

$$\log I_s = \log I_d + B \simeq \log I_d + 0.6. \qquad (3)$$

The ganglion cell's sensitivity to the two spots presented simultaneously was increased by 0.6 log units over that for either spot alone, while Ricco's law predicts that it should have increased by only 0.3 log units for a doubling of area; and, as described above, if the area of a spot is doubled by a change of diameter, Ricco's law holds. Doubling the area of a spot can be considered as the addition to the spot of a surrounding contiguous annulus of the same total area as the original spot. The question is why there is this difference between adding the additional area as a separate spot rather than as a contiguous annulus. Easter asserted that either summation of two disparate spots is somehow facilitated, facilitation increasing with separation, or summation is less complete for the annulus because of lateral inhibitory influences; clearly, the latter seems the more likely explanation.

Easter used the results of the two spot experiments to examine retinal function further. He assumed that the light in each spot is transformed into a quantity he called the excitation which is a function of intensity: e(I). Because of the spatial separation of the spots he assumed no lateral interactions[g] between them, and made the parsimonious assumption that excitations sum linearly. Thus the excitation when both are illuminated is twice the excitation from each. For equal responses:

$$e(I_s) = 2 \cdot e(I_d) \qquad (4)$$

Since Easter always compared the intensities needed to elicit the same response (sensitivity measure), any further processing after summation would not enter into the results. (For a more complete discussion of this aspect of sensitivity, see "Method of Sensitivity-Summation," below.) He then proposed that equations (3) and (4) can be reconciled by assuming a power law for e(I):

$$e(I) = C \cdot I^n \qquad (5)$$

from (3): $I_s = 10^B \cdot I_d$

so: $C \cdot (10^B \cdot I_d)^n = 2C \cdot (I_d)^n$

$$10^{Bn} = 2$$

$$n = \log(2)/B \simeq 0.5 \qquad (6)$$

Thus, Easter depicted a square root function on each pathway (possibly at the receptors), followed by linear summation. In the investigation of Ricco's law in which spot areas were changed, Easter postulated that just enough lateral interaction was present in the larger spot to cancel the advantage of spreading the light across a wider area.

But the necessity of postulating lateral interaction depends on the assumption that the sensitivity is uniform over the area for which Ricco's law holds. Easter (1968a) made this assertion, and illustrated it in his figure 1a, p. 256. In that figure, he showed log threshold sensitivity as a function of retinal position along two orthogonal axes; for the area in which he tested Ricco's law, he indicated a horizontal line to fit the data. Close examination reveals that for points along one of the axes the fit is somewhat tenuous; along the other axis it is still less believable. Later in the same paper, he presented another unit for which he plotted sensitivity as a function of position (his figure 4, p. 267). In this unit, it is clear that there is no significant plateau. And in his succeeding paper (1968b), a similar plot is given for a "red-off" unit (his figure 1, p. 275) and

again there is essentially no plateau. The data from these three units have been replotted in Fig. 3 on log-log coordinates. There is some scatter and variability, but it is not unreasonable to fit the data for radii greater than about 0.1 mm with a straight line having a slope of about -2. That is:

$$\log S \simeq -2 \cdot \log r + C'$$

or

$$S \simeq C/r^2 \qquad (7)$$

where S is sensitivity to a test spot of fixed area, C is a constant, and r is radial position of the spot in the receptive field. Equation (7) may be interpreted to state that sensitivity across the center of the receptive field falls off as the inverse square of distance from the center.* Naturally, the inverse square rule must break down for

*A sensitivity shift of this sort might seem to imply that the receptors differed in sensitivity as a function of their positions, but this is not the only interpretation. From the definition of sensitivity[g], and equation (7):

$$S_r = 1/I_r = C/r^2$$

so

$$I_r = r^2/C. \qquad (8)$$

The intensity must be changed at different positions to give the same responses; let us say that the reason for the change is that response at radius r has been multiplied by some factor ρ, where ρ is radius dependent. This is equivalent to saying the sensitivity is lowered at eccentric positions because fewer receptors contribute to the response. Thus, comparing the response function at r with that of the center:

$$\rho e(I_r) = e(I_o)$$

from (8):

$$\rho e(r^2/C) = e(1/C) \qquad (9)$$

If we agree with Easter's conclusion that e(I) is approximately a square root function, $e(I) = \sqrt{I}$:

$$\rho\sqrt{r^2/C} = \sqrt{1/C}$$

$$\rho r\sqrt{1/C} = \sqrt{1/C}$$

$$\rho = 1/r \qquad (10)$$

In short, stating that the response of receptors in any given area is a function of the square root of intensity of light incident on that area and that sensitivity at any point depends on the inverse square of the distance of that point from the center is completely equivalent to saying that receptors respond as the square root of intensity, and that the response from receptors in any given area are multiplied by a factor inversely proportional to their distance from the center of the receptive field.

Fig. 3. Sensitivity profiles of three units reported by Easter, replotted here on log-log coordinates. A: Unit for which Easter demonstrated Ricco's Law (Easter, 1968a, p. 256). B: Profile of center of "red-on" unit (Easter, 1968a, p. 267). C: Profile of "red-off" unit (Easter, 1968b, p. 275).

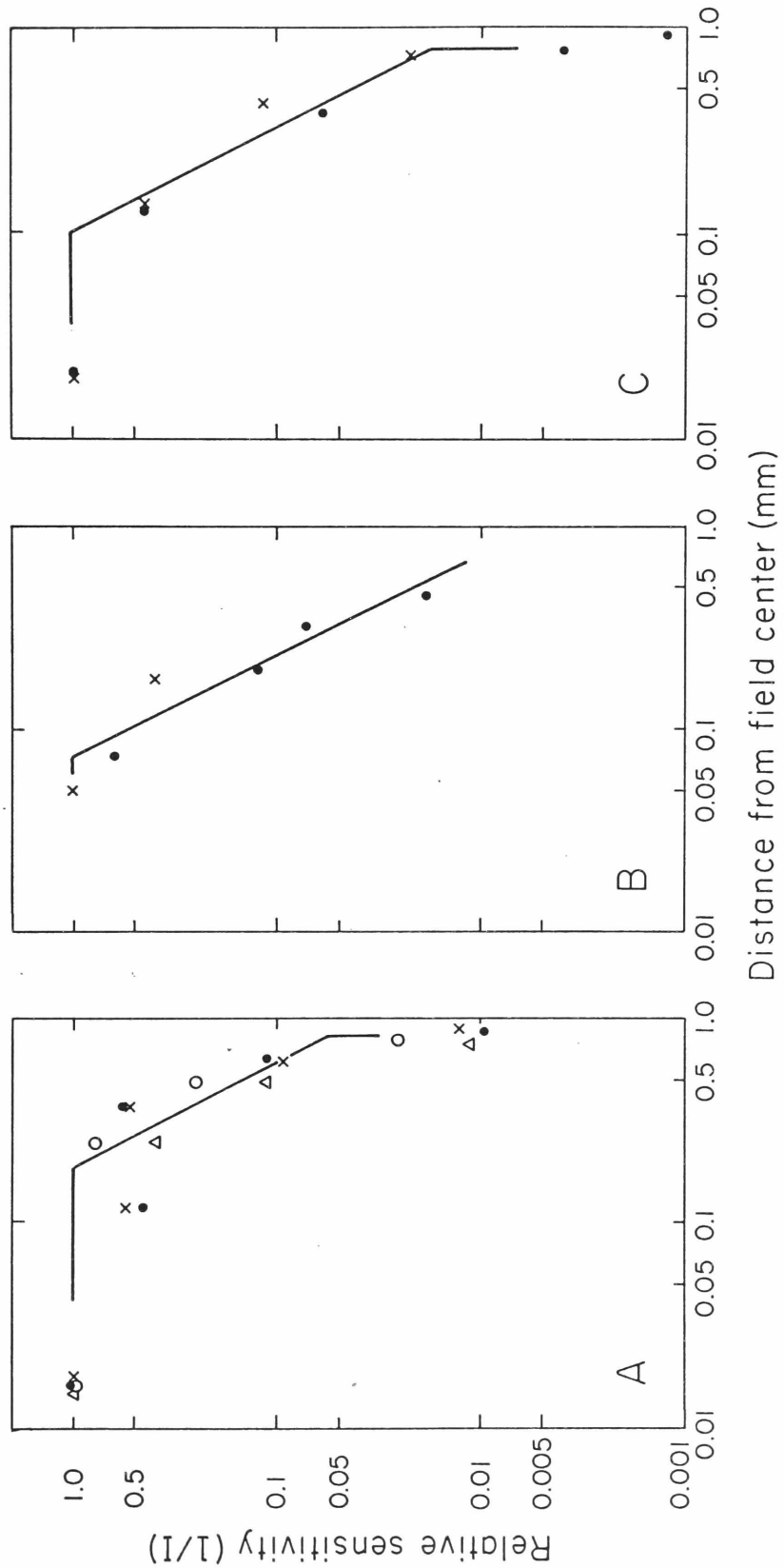
Relative sensitivity (1/I)
1.0
0.5
0.1
0.05
0.01
0.005
0.001
0.01
0.05
0.1
0.5
1.0
A
B
C
Distance from field center (mm)

Fig. 3

small radii, or we would be led to the absurdity of infinite sensitivity at an infinitesimal point in the center. Also, the center has only a limited extent, beyond which sensitivity rapidly falls to zero.

Let us now consider Ricco's law in the light of the above analysis. If the response in the central spot of area ΔA is e(I), then the response to an annulus at radius r, and width such that its area is also ΔA would be e(S · I), with S as given in equation (7). The net response R_{r_i} from a finite disk of radius r_i is the sum of the response contributions from each of the infinitesimal annuli of which it is composed:

$$R_{r_i} = \int_o^{r_i} e(S_i \cdot I_i)\,dA \tag{11}$$

Also $$dA = 2\pi r \cdot dr$$

so: $$R_{r_i} = 2\pi \int_o^{r_i} e(S_i \cdot I_i) r \cdot dr$$

For the same response to two disks of different arbitrary radii, r_1 and r_2:

$$R_{r_1} = R_{r_2}$$

$$2\pi \int_o^{r_1} e(S_1 \cdot I_1) r \cdot dr = 2\pi \int_o^{r_2} e(S_2 \cdot I_2) r \cdot dr. \tag{12}$$

And from (7):

$$\int_o^{r_1} e\left(\frac{C}{r^2} \cdot I_1\right) r \cdot dr = \int_o^{r_2} e\left(\frac{C}{r^2} \cdot I_2\right) r \cdot dr$$

Substituting the square root function of intensity for e(S · I):

$$e(S \cdot I) = \sqrt{S \cdot I}$$

$$\int_o^{r_1} \sqrt{(C/r^2) \cdot I_1}\; r \cdot dr = \int_o^{r_2} \sqrt{(C/r^2) \cdot I_2}\; r \cdot dr$$

$$\int_o^{r_1} \sqrt{I_1}\; dr = \int_o^{r_2} \sqrt{I_2}\; dr$$

The disks are uniformly illuminated, so I is independent of r; thus

$$\sqrt{I_1}\ r_1 = \sqrt{I_2} r_2$$

$$I_1 \cdot r_1^2 = I_2 \cdot r_2^2$$

or $$I_1 \cdot A_1 = I_2 \cdot A_2$$

for all A_1, A_2, which is simply Ricco's Law restated.

Thus, Ricco's law is completely consistent with a square root function for the receptors and an inverse square sensitivity function, both of which seem to obtain, without the assumption of any form of lateral interaction (except of course the simple summation of the diverse areas).* This is not to say there cannot be lateral interactions, only that they are not necessary to explain these data; moreover, if the data indicating an inverse square relationship were reliable enough to be completely irrefutable, one would be forced to say the lateral interactions must be independent of the separation of the interacting areas. Such an interaction, while it might be real in terms of what various cells are doing, would be mathematically undetectable from the point of view of the ganglion cell, being equivalent to some smaller multiplicative factor on the responses of the receptors.

Other workers have considered the summations of sub-areas within the center of the receptive fields of ganglion cells in the cat. Cleland and Enroth-Cugell (1968), recording from the cat, found that they could not observe the non-linearities reported by Easter (1968a) for the goldfish. They also measured the sensitivity profile of the center of the receptive field, and showed quite convincingly that there is a central plateau of constant sensitivity. Within this plateau, Ricco's

---

*Easter has also reached this conclusion, based on similar reasoning (personal communication).

law was obtained, for threshold sensitivities. Furthermore, when the relationships among separate stimulus spots within the center of the field were examined, it was found that sensitivities summed linearly, at threshold. That is, the points obtained by these workers would have obeyed the relationship given by equation (2), implying a linear system. It is important to note that these determinations were made at threshold, the lowest possible effective light level. There is evidence that the retina acts more linearly at low light levels: I shall present considerable evidence for such an effect in the course of this thesis; Easter noted it in his work, pointing out that "...there was a systematic tendency of the points at lower intensities to lie below [the line of unity slope and intercept greater than log 2 on log I_s vs log I_d coordinates], the ones at lowest intensity lying farthest below...."* Also in the cat, Creutzfeldt et al. (1970) examined Ricco's law at suprathreshold levels; the relationship was closely approximated at low light levels, but did not obtain at higher levels. When relatively large criteria were chosen, summation was more efficient for the larger areas, which is consistent with a non-linear receptor function such as that reported by Easter. It seems that in the cat sensitivity does not diminish with radial position in the compensatory fashion observed in the goldfish, so Ricco's law fails at levels high enough for the receptor non-linearity to be expressed. Easter's experiment was also repeated at suprathreshold levels in cat by Büttner and Grüsser (1968) with results nearly identical to Easter's.

Enroth-Cugell and Robson (1966) presented patterns of spatial sinusoidal gratings in cat receptive fields and found a magnitude and sign of response consistent with linear summation of the receptors in the field; it should be noted, however, that the maximum contrast available was only about 0.42 log units. Also, their experiment was a null-experiment, finding positions of gratings for which there was no response, and therefore makes no statement about any processing after the summation.

---

*Easter (1968a) p. 261.

Stone and Fabian (1968) investigated the linearity of summation of responses to two small stimulus spots within the center of a receptive field; using a wide range of luminances they compared the responses to both spots (physiological sum) to the arithmetically obtained sums of the responses to each spot presented separately (algebraic sum). Representing the response when both spots are illuminated as R_{1+2}, and the responses for each spot individually illuminated as R_1 and R_2, their criterion for linearity was

$$R_{1+2} = R_1 + R_2 \qquad (13)$$

Notice that these are responses compared to responses; any consistent non-linearity (e.g., at the receptors) which occurs before the first interaction of the sub-areas appears on both sides of the equation, and therefore cancels. The criterion will be met only if all processing from the first interaction to the recorded signal is linear; it will not be affected by non-linearities in each pathway before the first interaction (see Appendix II).

Stone and Fabian presented a plot of $(R_{1+2} - R_1)$ as a function of R_1 (parametric in R_2). For linearity, $(R_{1+2} - R_1)$ should always equal R_2; in fact, this was not the case: $(R_{1+2} - R_1)$ decreased steadily as a function of R_1. Thus, the criterion for linearity was not met, demonstrating a non-linearity at or after the first interaction of the two sub-areas.

Grüsser, Schaible, and Vierkant-Glathe (1970) also compared algebraic and physiological sums to test for linearity within the center of receptive fields of ganglion cells in cat. Rather than vary luminances, as Stone and Fabian did, they varied the number of spots which could be illuminated. They plotted directly the algebraic sum against the physiological sum, and found the algebraic sum always greater than the physiological; that is, rather than plotting a straight line of unity slope passing through the origin, they obtained a decelerated curve indicating a considerable non-linearity.

The majority of the data I have discussed so far were obtained from the cat and from the goldfish, two clearly dissimilar animals. Despite the difference in class, the results reported for cat have not been markedly inconsonant with those for the goldfish; there is hope of finding general principles that will be common to all vertebrates (see Dowling, 1970).

The data I have discussed do not as yet elucidate how the responses of the receptors are further processed by the retina. There is a non-linear transduction from light energy to receptor responses and responses from receptors in various areas are combined, but there appear to be further non-linearities in the retinal processing; there may also be other levels of interaction between sub-areas. We shall now turn our attention specifically to these possibilities. In particular, we shall be dealing only with interactions of receptors in different sub-areas of the center of the receptive fields of ganglion cells in goldfish. There are also chromatic interactions in goldfish; these I shall assume to have eliminated in the experiments to be described here by stimulating with near infrared lights, thus isolating the cone system containing only the "long" pigment. Chromatic isolation is further enhanced by using relatively low retinal temperatures (13° C) and maintaining the retina with pure O_2, both of which seem to depress the sensitivity of the "short" and "medium" cone systems (see "The Goldfish Retina," above).

Model of Retinal Processing

We shall now develop a model for the processing done by the goldfish retina. This model is intended to be general, allowing for a wide variety of the mechanisms and processes that seem plausible from previous work; it will serve as a framework for further discussion of retinal processing. Let us consider the output at the ganglion cell when two separate areas in its receptive field are stimulated.

The model for the center of the receptive field is depicted in Fig. 4. The receptors stimulated in each of the two separate areas (E) are treated

Fig. 4. General model of forms of processing postulated for the center of the receptive field of a goldfish ganglion cell. Lights are focused on two separate areas within the center of the field (I_1 and I_2); receptors in each area (E) absorb light and generate excitations (x_1 and x_2). The two excitations may interact laterally through the pathways marked c; the dashed line indicates that the interactions may be pooled (inclusive, so that the interaction on each pathway is dependent on all areas illuminated) or they may be specific (the effect on each pathway is due only to the activity in the other). The interaction might be either additive or multiplicative; it could be inhibitory or facilitatory; it might be non-recurrent (as shown) or recurrent. The blank boxes after interactions represent thresholds; the circle on the right shows the relationship between the output of such a box (ordinate) and the input to it (abscissa). The results of all of the preceding processing are two excitations (y_1 and y_2) which are summed. The box marked F represents any further processing of the sum, s, to give the nervous activity (spikes) recorded as the response, R.

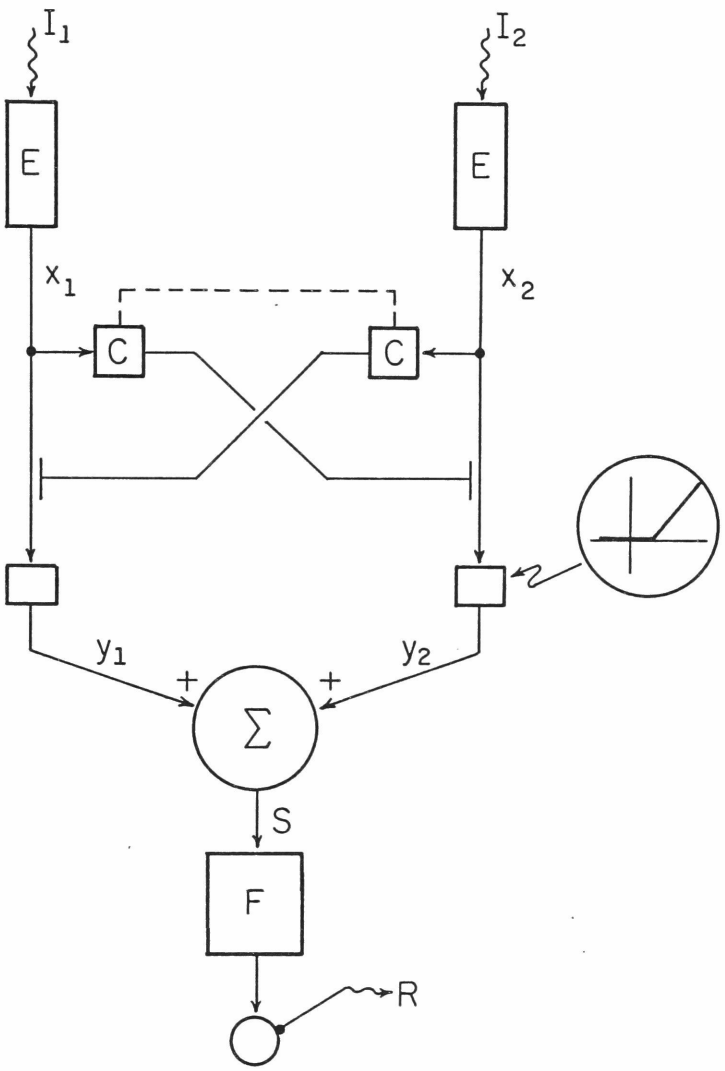

Fig. 4

as a single massed unit. Receptors effect a non-linear transformation in converting from illumination* to an output signal, x:

$$x = e(I) \tag{14}$$

It may be assumed that this function is the same for all cones containing a given photopigment, and that sensitivity differences between the two areas may be represented by a simple ratio multiplier ρ on one of the areas. ρ may be thought of as the number of receptors from one area relative to the other which contribute to the ganglion cell's response.

Ignoring for the moment the possibility of further distal processing, the outputs of the two areas must combine; each area has an effect on the ganglion cell. I shall assume that this combination is linear, at least between sub-areas both in the center of the receptive field or both in the surround. Linear summation is consistent with the findings discussed in the previous section, and most forms of non-linear combination may probably be modeled as linear summation preceded or followed by a non-linearity. This need not be true for the combination of antagonistic influences (e.g., center and surround or spectrally opponent inputs) which might be divisive (Sperling, 1970).

Let us now consider the possibility of lateral interactions[g], as distinguished from the spatial summation described in the preceding paragraph. Lateral interaction, indicated in Fig. 4 by the pathways labeled "c," will be taken to mean an influence whereby light in one area may affect the response to light in another area; it is assumed to involve a distal retinal pathway, and therefore occurs before spatial summation. Lateral interactions may be linear (subtractive) or non-linear

---

*Only light that is absorbed by the photopigment in a receptor can lead to excitation; thus, the effectiveness of illumination in stimulating a receptor depends on the product of the light's power spectrum and the absorption spectrum of the photopigment. In the experiments described in this thesis, the problem is simplified by always using monochromatic light whose wavelength is such (near infrared) that it is absorbed only by cones containing the "long" pigment.

(e.g., shunting[g]); they may be specific, in that the effect on one area is a function only of other areas, or they may be inclusive, in that the interactions depend on all illuminated areas (Furman, 1965; Grüsser et al., 1970). Another important distinction is between recurrent and non-recurrent interactions. In a recurrent system, the interactions depend on the outputs of the interacting elements; in a non-recurrent system, interactions depend on the inputs. If, as in Fig. 4, the receptor outputs are designated x_i, and the inputs to the summation are y_i (after interactions), a non-recurrent system would have the relationships:

$$y_i = \Phi(x_i, \sum_j c_{ij} x_j)$$

where Φ is a function depending on the form of interaction, and c_{ij} is the effect of element j on element i. (The system in Fig. 4 is non-recurrent as drawn.) A recurrent system, on the other hand, would have the relationships:

$$y_i = \Phi(x_i, \sum_j c_{ij} y_j)$$

If one observes only the steady-state responses of the ganglion cell the two forms are indistinguishable, but there may be differences that could be seen in the transient responses.

There are two additional properties one might expect of lateral interactions, though neither is necessary: 1) The strength of interaction should depend on the spatial separation of the areas. In general, one would expect interactions to diminish with increased separation. 2) A distal interaction plexus should probably not be cognizant of receptive field boundaries, since two areas both within the center of the receptive field of one ganglion cell could lie one in the center and one in the surround of the receptive field of some other ganglion cell.

Most forms of lateral interaction may be roughly approximated (over a short range) by subtractive interactions, such as are found in the horseshoe crab, Limulus (Ratliff, 1965). But a strictly linear interaction system would be undetectable at the ganglion cell, barring temporal distinctions. For a linear non-recurrent interaction, the

output, y_i, at each point would be (assuming an inhibitory interaction):

$$y_i = x_i - \sum_{j=1}^{\infty} c_{ij} x_j \qquad (j \neq i)$$

The total signal after areal summation, s, is thus:

$$s = \sum_{i=1}^{\infty} y_i = \sum_{i=1}^{\infty} [x_i - \sum_{j=1}^{\infty} c_{ij} x_j] \qquad (i \neq j) \qquad (15)$$

Equation (15) may be rewritten as:

$$s = \sum_{j=1}^{\infty} [x_j - \sum_{i=1}^{\infty} c_{ij} x_j] \qquad (i \neq j)$$

or

$$s = \sum_{j=1}^{\infty} x_j (1 - \sum_{i=1}^{\infty} c_{ij}) \qquad (i \neq j) \qquad (16)$$

The sum on i in equation (16) is a constant for each j; this equation simply expresses the sum of all active areas (x_j), with an attenuation factor ($1 - \sum_{i=1}^{\infty} c_{ij}$). It is mathematically equivalent to summing independent areas.

Equation (16) results only because both positive and negative signals are tolerated if the interactions are completely linear. In Fig. 4 for example, if only area 1 were illuminated, $y_1 = e(I_1)$, but $y_2 = -c \cdot e(I_1)$; the pathway from area 2 provides a negative input to the areal summation point. We might postulate that an excitatory area can only provide excitatory signals (lacking an inhibitory transmitter, for instance); if the signal at y_2 is negative, it would have no influence. This is a special case of a threshold (threshold = 0), as shown in the circled inset in Fig. 4. In general:

$$y = \begin{cases} x - x_o & ; x > x_o \\ 0 & ; x \leq x_o \end{cases} \qquad (17)$$

where x_o is the threshold value of x). If a threshold is included in each pathway, as shown in the figure, equation (16) no longer applies, and the interaction is effectively made non-linear. Then, if area 1 is stimulated alone, the input to the summation (s_1) is:

$$s_1 = y_1 = x_1 - x_o \qquad (18)$$

(Because $x_2 = 0$, $x_2 - cx_1 < x_o$ and $y_2 = 0$.) But if area 2 is stimulated at the same time as area 1, the influence of area 1 will be:

$$s_1 = x_1 + (-cx_1) - x_o$$

$$s_1 = x_1(1 - c) - x_o \qquad (19)$$

$(x_2 - c \cdot x_1 > x_o)$ The difference between equations (18) and (19) makes the interaction detectable at the ganglion cell level; it expresses a non-linearity.

In the absence of any lateral interactions, the thresholds themselves are a non-linearity if $x_o > 0$. Equation (17) is approximately linear if $x >> x_o$; as $x \to x_o$, the equation becomes non-linear, until the sharp discontinuity at $x = x_o$ and $y = 0$.

Another possible non-linearity could exist after the combination of the responses of the two areas. It is quite reasonable that there might be an additional non-linear summation, or a compressive[g] function such as the one at the receptors.

This general model incorporates the features suggested by the work reviewed earlier in this chapter. It allows for a non-linearity at the receptors (e.g. Tomita, 1970; Boynton and Whitten, 1970), lateral interactions (e.g. Easter, 1968a; Grüsser et al., 1970), linear summation of sub-areas (e.g. Enroth-Cugell and Robson, 1966), and non-linearity after summation (e.g. Spekreijse and van den Berg, 1971; Stone and Fabian, 1968). We must now turn to direct experimentation to see which, if any, of its features can be confirmed or denied.

MATERIALS AND METHODS

Biological Preparation

The subjects of these experiments were common goldfish (Carassius auratus) from 15 to 22 cm in length, obtained from Ozark Fisheries, Inc., Stoutland, Missouri. The fish were stored in large tanks at 22° C, and fed daily on commercial fish food. While in storage, the fish were maintained in a cycle of 12 hours of light, 12 hours of darkness.

About 2 to 4 hours before surgery, the fish were removed from the storage tank and placed in a small light-tight tank, at room temperature. This was to allow the photomechanical processes which occur in goldfish during dark adaptation to take place (Walls, 1942). Thus, at the time of surgery, the processes from the pigment epithelium which surround the receptors in light adaptation were fully retracted, the rod myoids were fully contracted, and the cone myoids fully elongated. This state, which was confirmed by examination of histological specimens, is optimal for the mechanical separation of the retina and pigment epithelium with minimal damage to the receptor outer segments.

The surgical procedure used was essentially the same as that described by MacNichol and Svaetichin (1958). The entire procedure, which could be performed under normal room lights due to the sluggishness of the photomechanical processes, required about 3 1/2 minutes. Using a curved spatula and curved scissors, the eye was removed from the fish, which could then be returned to water until the second eye was required. Next, the anterior portion of the eye was removed. The resulting eyecup was tipped forward, and the edge of the retina grasped with a fine forceps and teased forward out of the cup, leaving behind the pigment epithelium and many of the rod outer segments. When the retina was about half out of the eye, microdissecting scissors could be inserted behind the optic disc, between retina and pigment epithelium, and the optic nerve severed. The optic nerve fibers, which had been under slight tension before cutting, apparently withdrew through the disc, leaving a small hole in

the retina. The retina was then free to be pulled entirely clear of the eye, completing the operation.

The retina, receptor outer segments uppermost, was placed on a glass slide in the bottom of a moist chamber. The vitreous humor of the goldfish is quite viscid, so that the retina rested on a pad of humor. This provided the mechanical support to keep the retina in a slightly curved position in the chamber. In the chamber, light was directed onto the retina from below, while the recording electrode was advanced from above.

The chamber was made from brass, with a glass floor to permit light to be projected onto the retina from below. The lower part of the chamber was thick, with a duct system for circulating cooling water from a Lauda K2/R Cooler (Brinkmann Instruments); the temperature of the retina was maintained at 13^{o} C. The temperature was measured by a tele-thermometer (Yellow Springs Instrument Co., model 426C tele-thermometer with a #9545-C15 temperature sensing probe); the sensing probe was placed in a pad of cotton soaked in fresh-water teleost Ringer (Cavanaugh, 1956) on the glass plate where the retina normally is placed.

The upper part of the chamber also had a duct system, through which a steady flow of moist oxygen from a compressed gas tank was directed at the retina from each side of the chamber. The flow was maintained at 90 ml/min by pressure reducing and needle valves, and monitored on a flowmeter (Matheson Gas Products, #622PBX) that had been calibrated by timing the rate at which gas from its exit displaced water. The oxygen was bubbled through distilled water for humidification, and passed through glass wool to remove any suspended water droplets before entering the chamber. In the winter, when relative humidity was low, additional moisture was provided by placing wet cotton balls in the corners of the chamber. The top of the chamber was only partly covered, providing access for the recording electrodes, and egress for the excess oxygen and waste gases; the latter process was assisted by a vacuum line placed near the top of the chamber. The partly enclosed volume of the chamber in which the retina was placed was 4.6 x 4.6 x 1.2 cm in size.

Electrical Recording

Recordings from single ganglion cells were made by penetrating the retina to the depth of the ganglion cell layer with fine electrodes of platinum-iridium wire. The method for manufacturing these electrodes was the same as that given by Snodderly (1969), derived from Wolbarsht and Wagner (1963). Short lengths of the wire (10 mil, 70/30 platinum-iridium), in holders made from modified innoculating needle holders, were suspended vertically in a solution of 6M NaCN in 30% NaOH. Alternating current was passed between the wires and a carbon electrode in the bottom of the solution, etching the wires to a long tapered point less than 1μ at the tip.

The electrodes were insulated by impaling a bead of molten solder glass which was suspended on a U-shaped loop of platinum wire and kept molten by passing electric current through the wire; when a sufficient length of shank had been coated, the electrode was removed laterally through the open end of the loop. The glass was then rendered hydrophobic by dipping the electrode in hexamethyldisilazane. Before using the electrode, the exposed tip was coated with platinum black by placing the electrode in 0.1% H_2PtCl_6 solution and passing about 1 μa electric current from an indifferent wire. The platinum blacking was repeated before each use of the electrode.*

The indifferent electrode was a silver-silver chloride wire in a glass tube of fresh-water teleost Ringer solution (Cavanaugh, 1956), with a cotton wick contacting the edge of the vitreous pad. The active electrode was clamped in a micromanipulator (designed and built by The Rockefeller University Instrument Shop) which poised it vertically over the retina.

---

*Electrodes could often be used repeatedly, the greatest hazard being the accumulation of a mass of tissue on the tip (despite the hexamethyldisilazane). In this eventuality, the electrode could be restored by oxydizing the offending material in the molten glass bead, and, as this melted the insulation recoating it as if it were newly etched.

The voltage difference between the active and indifferent electrodes was amplified by a small battery powered capacitively-coupled pre-amplifier with a field-effect transistor input (designed by M. Rossetto of The Rockefeller University Electronics Shop). The moist chamber and the pre-amplifier were enclosed in a large steel box, which shielded against stray light and electromagnetic signals. The output of the pre-amplifier was further amplified, and then filtered by a Krohn-Hite adjustable filter (Model 3103). The filtered signal was the input to a level detector (designed by M. Rossetto); this device provided a shaped output pulse whenever its input signal exceeded an adjustable threshold level. The pulses were recorded on a 7 track tape recorder (Honeywell, Model 8100), along with several channels of frequency encoded information defining the stimulus regime. One of these channels successively encoded the wavelength in each beam of the optical system (see "Optical Stimulator," below), intensity in each beam, and an automatically incremented protocol number by which the record could be identified. Another channel encoded the current state of the electromagnetic shutters in each beam: which ones were open and which were closed. A pulse code was used to avoid the relatively large memory requirements for encoding an analog signal when the tape was later played into a Control Data Corporation 160A digital computer. The computer generated BCD tapes for further analysis and permanent records.

The experiments were controlled by a bank of digital timers (designed by L. Eisenberg of The Rockefeller University Electronics Shop) which could be programmed to open and close shutters and start and stop the tape recorder. A bank of counters sensing the pulses from the level detector could be used to display the number of action potentials fired in the second before a shutter opened ("spontaneous"), during the second the shutter was open ("on" response), and the second after it closed ("off" response).

Optical Stimulator

One of the advantages of the isolated retina preparation is the relative ease with which optical stimuli can be controlled. With no wavelength selective tapetum or pigment epithelium behind the retina to reflect light back through it, chromatic control is considerably simplified. With the electrode and light entering from opposite sides of the retina there is no shadow cast by the electrode.

The system used in these experiments consisted of three identical beams whose outputs were superimposed on the retina. The stimulus size, shape, position, wavelength, quantal flux, and timing were independently controllable in each beam. A typical beam is shown in Fig. 5.

The light source in each beam consisted of a tungsten filament quartz-iodide lamp (50 W, 60 V/T3 1/2-in prefocused base, Sylvania Electric Products), driven by a programmable power supply. The lamp filament was imaged on the entrance slit (2 mm) of a grating monochromator (Jarrell-Ash, 1/4 Meter Ebert Monochromator); the band width was about 8 nm. Immediately past the exit slit (2 mm) was a microscope cover slip which diverted a small percentage of the light from the beam to a silicon photocell.

The output of the photocell was used as a negative feedback on the lamp's power supply, closing a servo loop (for a more complete description, see Rosen, Levine, Rossetto, and Abramov, 1970). Since the photocell was located after the monochromator, its spectral sensitivity would have determined the spectrum of the light beam had the photocell's response been used directly as the feedback signal. To overcome this, a Vernistat Variable Function Generator (Perkin-Elmer function adjusting assembly model DC250, and interpolating potentiometer model 2X5) provided a variable gain in the feedback loop, the gain being an adjustable function of the wavelength setting of the monochromator. The functions in each beam were adjusted such that each delivered to the retina an equal quantal flux spectrum. This system not only provided a means of achieving the desired spectrum, but

Fig. 5. Diagram of a typical beam of the optical stimulator. Light from the lamp is focused on the monochromator, whose output is sensed by the photocell, which in turn controls the lamp voltage. The monochromatic light passes through the aperture of an electromagnetic shutter and two wheels of neutral density filters before illuminating the field stop. The field stop is imaged at the level of the receptors in the retina. The three beams of the optical system are joined by the mixing cubes, which also provide an observation beam so the experimenter may observe the stimuli.

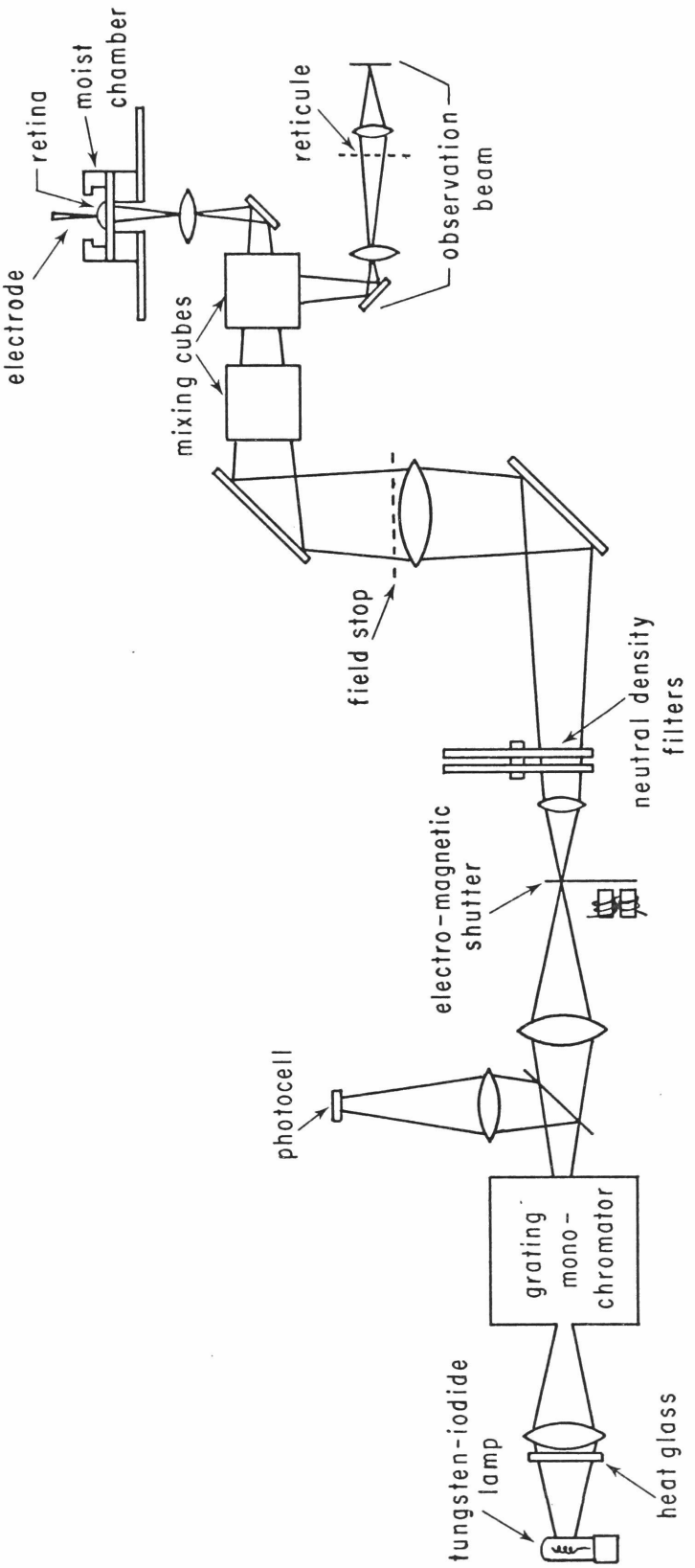
retina
moist chamber
reticule
observation beam
electrode
mixing cubes
field stop
neutral density filters
electro-magnetic shutter
photocell
grating mono-chromator
tungsten-iodide lamp
heat glass

Fig. 5

automatically compensated for aging effects of the bulb. Also, being a dynamic servo system, the light could be temporally modulated. A sinusoidal signal could be added to the input of the servo loop: as the feedback photocell monitored the light output of the monochromator, the current through the bulb was adjusted such that the light followed the appropriate sine wave, up to the limit of the bulb's ability to cool rapidly enough.

The output of the monochromator fell on another lens, which formed a real image of the lamp filament on an aperture. At this aperture, where all the light is channeled through a small area, the vane of a shutter can interrupt the light with minimal motion. The electromagnetic shutter used (see Hartline and McDonald, 1947) had a rise and fall time of less than a millisecond. The shutter vane also covered a second aperture, through which light from a neon bulb fell on a photocell; this photocell provided a record of when the shutter actually opened or closed.

The light passing through the shutter aperture was collected by a lens and passed through two wheels of Inconel neutral density filters (Boxton-Beel Inc.), with which it was possible to change the relative quantal flux over a range of 4.8 log units, in 0.2 log unit steps. The collecting lens also served the purpose of projecting a real image of the monochromator grating; this image is of special significance, as it is at this point that the cross section of the beam is most uniform in luminance and chromaticity. Hence, it was at that image plane that the field stops, which determined the stimulus shape, were placed. The field stops were made photographically, with about 6 log units difference in density between clear areas and mask. A large selection of field stops had been made, providing a wide range of possible spots, annuli, and edges. Also at the field plane, and therefore having practically no effect on the grating image, was a large field lens, whose function was to reimage the filament on the final projection lens.

The three beams were joined by mixing cubes and passed down a common path to the final projection lens. This lens served to image the three field planes (reduced 10:1) through the glass bottom of the moist chamber and onto the retina. The images of the field stops were thus the stimulus pattern on the receptors, entering in the normal physiological direction. The final mixing cube also provided an observation beam which could be viewed by the experimenter. In this beam, the field stops were imaged on a reticule which was marked with a grid whose scale corresponded to millimeters on the retina; this provided a convenient means of positioning the field stops. An opaque block in the path to the retina could be used to avoid illuminating the preparation during alignment of the field stops.

The spectral characteristics of each beam of the optical system were calibrated with a highly linear silicon photocell placed in the plane of the retina; the output of the photocell was read on a sensitive current meter. The spectral characteristics of this photocell were calibrated by Dr. E. F. MacNichol, Jr., of the Johns Hopkins University Biophysics Department. Absolute sensitivity of the photocell was obtained by calibrating it against a radiometer (Yellow Springs Instrument Co., model 65). The maximum flux density (0.0 log units) from each beam was 2.75×10^{12} quanta$\cdot$cm$^{-2}\cdot$sec^{-1}.

The neutral density filters were individually calibrated by the manufacturer, using a Cary 14 spectroradiometer. Additivity of the filters was checked by the photocell in the plane of the retina.

The scattered light in the system was reduced by adding to the final projection lens an aperture which just passed the lamp's filament image. With the aperture in place, flux density in the center of an annulus with a 4 mm outer diameter and 2 mm inner diameter was less than 1% of the annulus flux density. Scatter flux density from a 1/2 mm spot was less than 0.2% of the spot flux density, measured at the retina.

Experimental Procedure

After the retina was removed and placed into the chamber, it was left in total darkness for about 10 minutes. This allowed some time for the retina to settle mechanically, and reach thermal equilibrium.

When an electrode was being prepared for a penetration, a small spot was placed in one of the beams and aligned with the reticule center. The spot could be observed on the retina with a dissecting microscope, and used to position the electrode tip. Also, the final projection lens could be adjusted slightly to focus the spot sharply at the level of the outer segments of the receptors. During the search for a single unit, the retina was stimulated with a one second flash of light every four seconds. This "searching" flash was necessary to insure that units normally silent in the dark would not be overlooked. If no unit was found, the electrode was withdrawn and, rather than move the electrode (which was concentric with the optical system), the moist chamber (and hence retina) was shifted slightly before the next penetration.

When a unit was isolated, the "searching" light was extinguished, and the receptive field of the unit examined by a series of one second test stimuli in order to classify the unit. Units that responded at "off" to a spot (710 nm) in the center of the receptive field and "on" to an annulus (710 nm) in the surround were classified (-R+R); units with the opposite responses to deep red stimuli were classified (+R-R). Thereafter, stimuli of one second duration were presented strictly one every 25 seconds; this regimen was maintained for several minutes before recording was begun, and continued uninterrupted (as far as was feasible) throughout the experiment. (The only exception was the test flash in the relative timing experiment described below.) The cell's activity was recorded for a five second period beginning one second before each stimulus. This five second interval I shall refer to as a "record."

The only stimuli used in this work were lights of 710 nm wavelength. Other wavelengths were used in searching for units and in informally classifying the cell types, but the actual experimentation

was all performed with this near infrared stimulus.

Experimental Paradigms

The experiments described in this thesis fall into two basic categories. The first, which I call "relative timing," was a simple test for non-linearities other than those immediately at the receptors; the second, which I call "two spot summation" (though the stimuli were not always spots) was intended to provide a more complete basis for analyzing the retina. The two spot summation experiment actually consisted of two different, but related, experiments.

Relative Timing The relative timing experiment tested the effects of a "conditioning" stimulus (1 sec) on the response to a brief (200 msec) "test" stimulus. The conditioning stimulus was always placed on the receptive field such that the cell's response to it alone was an increased "on" firing; e.g., in the surround of a (-R+R) unit. The test flash, which was on a separate area of the receptive field from the conditioning stimulus, might give either an "on" or "off" response.

Each record consisted of one presentation of the conditioning stimulus and the test stimulus. The relative time of occurrence of the test stimulus was varied from record to record; in some records it preceded the conditioning stimulus, in some the two began simultaneously, and in some the test fell entirely within the time in which the conditioning stimulus was on. There were also a small number of "control" records in which either the test or the conditioning stimulus was omitted. A typical series began with the onset of the test flash preceding the onset of the conditioning stimulus by 400 msec; the difference in relative timing was changed by 100 msec from record to record. The entire series was then repeated, starting at 350 msec. Thus, points were obtained every 50 msec, with at least three of each control record taken at various points in the experiment. The entire experiment could be completed in about 10 minutes.

Two Spot Summation: General In the two spot summation experiments, two spatially separate loci of equal area were stimulated, either both within the center of the receptive field or both within the surround. These stimuli were presented either simultaneously or one at a time, over a wide range of intensities--this yielded intensity-response functions for each area alone and for the two together. Usually about 6 to 8 different intensities were used, spanning the range from threshold to saturation. The order of presentation of the stimuli was varied (i.e. whether one area, or both at once would be presented, and at which intensity), so that systematic effects due to precedent stimuli were minimized. Also, any consistent trend due to a long term change in the preparation was thereby randomized within the data. About 15 to 20 minutes were required to obtain a sufficient body of data for analysis. There were two forms of two spot experiments.

Two Spot Summation: Equal Intensities The first and more significant of the two spot summation experiments was the "equal-intensity experiment." In this experiment, for the records in which both areas were stimulated simultaneously, the two stimuli were of the same intensity. Since the stimuli were also of equal area, the same total flux was being delivered to each area. The intensity used was varied from record to record, covering the same range as was used in stimulating each of the spots alone.

Two Spot Summation: Peg Experiment The second of the two spot summation experiments I call the "peg experiment." In this case, when both areas were simultaneously stimulated the intensity in one of the areas was held (pegged) at a constant, fixed value while the intensity in the other area was varied over a wide range. Obviously, the peg experiment intersects the equal intensity experiment at the point at which the intensity in the area allowed variable intensities has the same value as the fixed intensity in the other area. Both areas were stimulated for one second, as in the equal intensity experiment, but the intensity in only one of the areas was changed from record to record.

Response Measure

The measure of a unit's response to a stimulus will always be expressed in adrians*; responses were taken as the average firing rate over a one second interval. (Some exceptions in the relative timing experiment and a special form of analysis of the two spot summation experiments will be discussed later.) The interval chosen was either the one second during the stimulus, or the second immediately following it, depending on the nature of the response. The excitatory portion was always the part used: this avoided the difficulty of analyzing responses which are limited by the impossibility of rates below 0 adrians. Thus, for the unit shown in Fig. 6, the excitatory part of the responses were measured in the second immediately following stimulus offset ("post" curve of the intensity-response curves on the upper right). Stimuli in the surround of this same unit's receptive field yielded excitatory "on" responses (lower portion of Fig. 6), and these responses were therefore measured over the one second during the stimulus ("during" intensity-response curve on the lower right).

The responses are expressed in absolute firing rates, with no correction for the cell's spontaneous firing in the dark. When applicable, the spontaneous rate is indicated in the figures by a solid line; it was obtained by averaging the firing rate of the unit in the one second periods preceding the stimuli whose responses are being shown.

---

*The adrian has recently been introduced as a measure of nerve firing rates (Shapley, 1971); one adrian is defined as one nerve impulse per second.

Fig. 6. Responses from the center and surround of a goldfish ganglion cell stimulated with deep red (710 nm) light. Top left: Responses to a 1 mm spot centered on the receptive field. The inset shows the oscilloscope tracing in response to an intensity of -1.0 log units relative to the maximum; the solid horizontal line indicates the stimulus duration (1 second). The topmost graph is firing rate vs time for the same record. Stimulus duration is indicated by a darkening of the abscissa and two vertical lines. Firing rate is plotted for successive 50 msec periods, according to the method described by Schoenfeld (1964). The second graph is the response to the same stimulus at -2.0 log units (i.e. one tenth the intensity used in the record above). Top right: Intensity-response curve for a 1 mm spot on the receptive field center (sketched at the top of the graph). The horizontal line indicates the spontaneous rate; circles and solid curve represent the average firing rates in the one second after the offset of the stimuli; triangles and dashed curve are the averaged firing rates during the stimuli. Averages derived from the two records on the left are shown as filled symbols. Bottom: Oscilloscope tracings, firing rate as a function of time, and intensity-response curves from the same unit in response to the large annulus indicated in the sketch. Conventions as above.

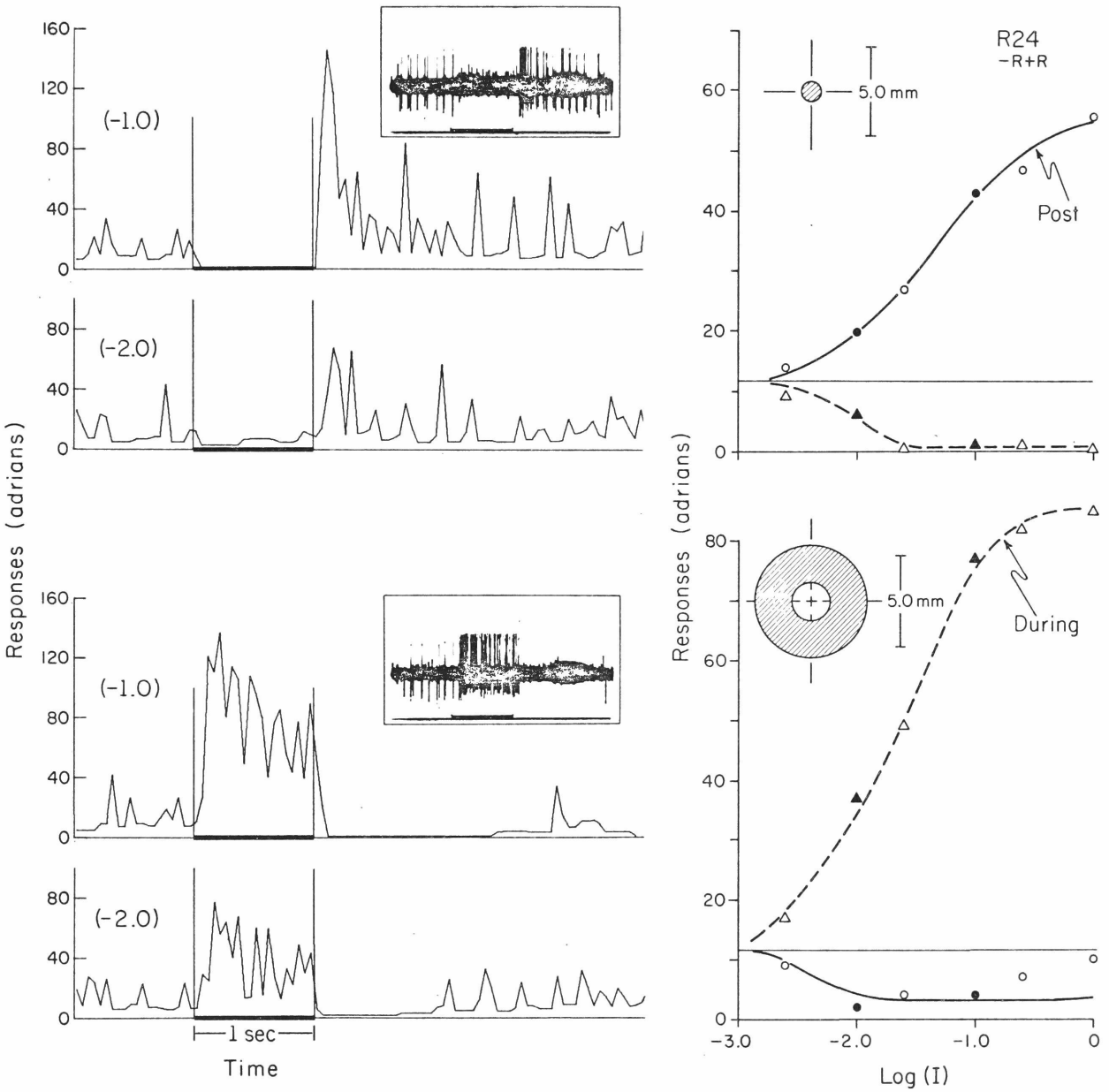
160
120
80
40
0
(-1.0)
(-2.0)
1 sec
Time
Responses (adrians)
R24
-R+R
5.0 mm
Post
During
60
20
-3.0
-2.0
-1.0
Log (I)

Fig. 6

METHODS OF ANALYSIS

The data acquired in the experiments described in the preceding section require further analysis and interpretation before they can aid in an understanding of the processing in the retina. The particular analyses by which I interpret the data are predicated upon the model of the receptive field center (or surround) outlined in Fig. 4.

In this section, I shall describe how I have analysed the data from the two spot experiments. My analysis of the relative timing experiments (see "Relative Timing" in Experimental Results and Conclusions) was neither unique nor especially revealing of the mechanisms in the model; it does not merit special attention in this section. My analyses of the two spot experiments, however, though a direct outgrowth of methods described in the Introduction, are somewhat novel and generally applicable. There are two analysis methods I shall consider; the two are complementary and much of the power of this analysis derives from the use of both on the same data.

The Method of Response-Summation

My first method for analysing the two spot experiments is fundamentally the same as that used by Stone and Fabian (1968), and results in the same coordinate system used by Grüsser *et al*. (1970). The method consists of comparing the arithmetically summed responses recorded when each spot is stimulated separately (i.e., $R_1 + R_2$) to the response of the ganglion cell when both are stimulated simultaneously with the same intensities as before (R_{1+2}, the physiological sum). If the system is linear from the point at which the two areas first interact until the output, the following linearity criterion must be met:

$$R_{1+2} = R_1 + R_2 \qquad \text{(linear)} \qquad (1)$$

Notice that this statement must be true regardless of the form of the non-linearities before the areas interact, even if the non-linear functions within the two areas are quite different. These non-linearities each appear on both sides of equation (1), and therefore cancel.

Equation (1) should be tested for as many different intensity conditions as is feasible; to present such data in a meaningful way, I plot them on a coordinate system such as that used by Grüsser et al:

$$\text{abscissa} = R_1 + R_2$$

$$\text{ordinate} = R_{1+2} \qquad (2)$$

Such a system I call "response-summation coordinates," and the graph thereon a "response-summation plot." Each point on a response-summation plot is a test of equation (1); summation is linear for the conditions represented by a point if it lies on the straight line of unity slope passing through the "spontaneous"[g] point (the point obtained when no light is used in either spot).

Fig. 7 demonstrates the means of generating a response-summation plot from the data collected in the two spot experiments. The particular unit from which these data come was "red-off" in the center, and had a spontaneous firing rate of 12.1 adrians. The stimuli are shown in the lower right of the figure; both spots were within the center of the receptive field, so the responses shown in this figure are all "off" responses.

The curves on the lower left coordinates are the individual intensity-response curves for each of the two areas illuminated separately. That the two curves do not superimpose indicates that the two areas are of unequal sensitivity, area 1 (triangles) being more sensitive than area 2 (squares). Each symbol represents the response to a single presentation of light; the curves were fit by eye. Notice that the intensity range of the experiment exceeded 2 log units.

The curves on the upper left coordinates are the intensity-response curves obtained when both spots are simultaneously illuminated. The dotted curve, fit by eye to the open circles, results from the equal-intensity experiment (see "Experimental Paradigms"). The solid curves, fit to the filled symbols, result from two separate peg experiments. (As far as is feasible, these conventions will be followed in all future presentations of data from two spot experiments.)

Fig. 7. Response-summation analysis of the responses of a goldfish ganglion cell. The numbers and letters in the top left corner of the figure are, respectively, the cell number (R52) and response type (-R+R) of this unit; the center was inhibited ("off" responses) by red light while the surround was excited ("on" responses). The stimuli, two spots of 710 nm light 0.4 mm in diameter tangent at the center of the receptive field, are sketched on the lower right. Lower left: Intensity-response curves for each area stimulated alone (area 1 is shown by triangles, area 2 by squares). Spontaneous firing rate of the unit is indicated by the horizontal solid line. The solid curves were fit to the data by eye. Upper left: Intensity-response curves for both areas stimulated simultaneously. Open circles and dotted curve indicate responses from the equal-intensity experiment; the solid symbols and curves are for responses from two peg experiments (the solid circles are for the intensity in area 1 held at -0.6 log units; the triangles for an intensity in area 1 of -2.0 log units). Upper right: Response-summation plot, R_{1+2} vs $(R_1 + R_2)$; symbols as on the upper left. The curves are derived from the curves on the intensity-response plots; the dotted curve is from the equal-intensity experiment, and the solid curves from the peg experiments. Spontaneous rate (zero response) shown by solid vertical and horizontal lines. The dashed lines between the graphs indicate the method of locating one particular point on the response-summation plot, at R_{1+2} = 42 adrians, $(R_1 + R_2)$ = 65 adrians, as described in the text.

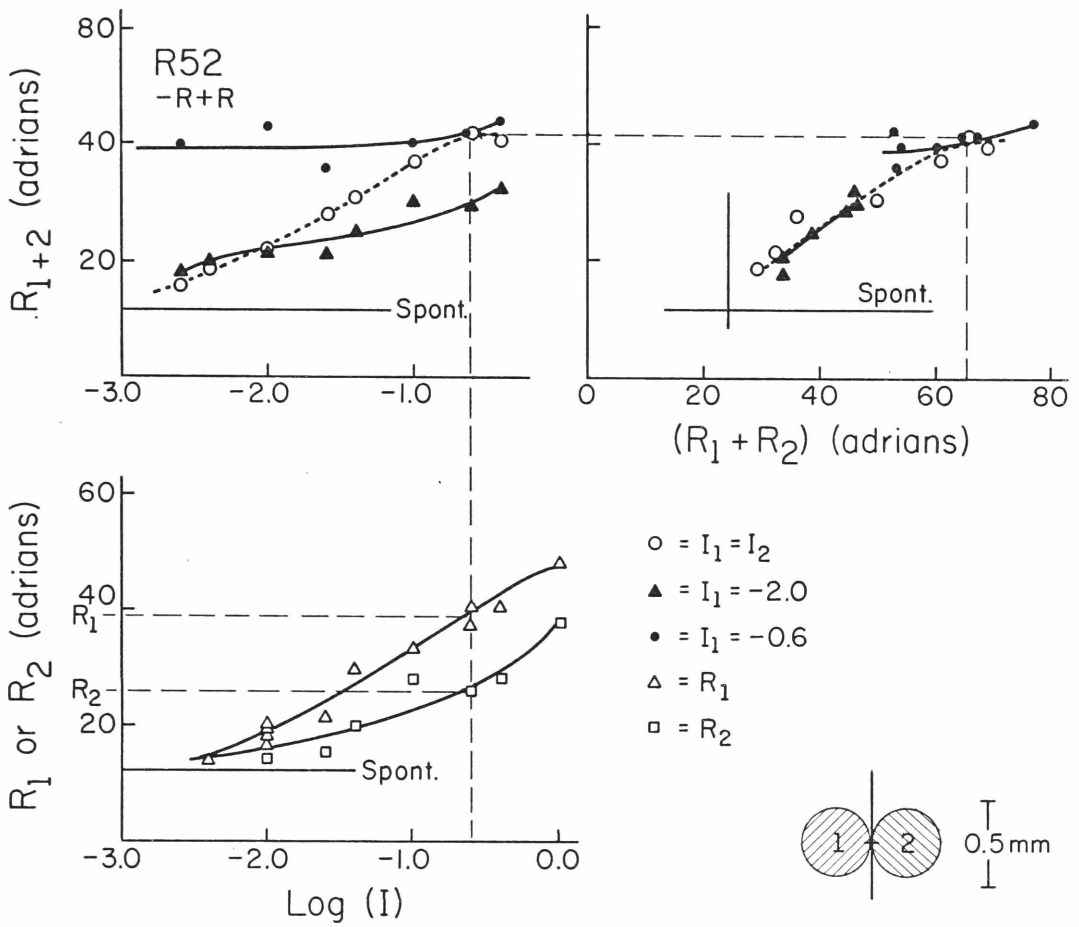
R52
-R+R
R_{1+2} (adrians)
80
40
20
Spont.
-3.0
-2.0
-1.0
0
20
40
60
80
$(R_1 + R_2)$ (adrians)
R_1 or R_2 (adrians)
60
40
20
R_1
R_2
Spont.
-3.0
-2.0
-1.0
0.0
Log (I)
○ = $I_1 = I_2$
▲ = $I_1 = -2.0$
● = $I_1 = -0.6$
△ = R_1
□ = R_2
1
2
0.5mm

Fig. 7

The response-summation coordinates, as defined by equations (2), appear on the upper right in the figure. The broken lines between the various graphs indicate the method of plotting one point (an open circle) of the response-summation plot, the point obtained from the equal-intensity experiment for the intensity = -0.6 log units. (Since -0.6 happens also to be the intensity at which area 1 was fixed in the peg experiment shown by the solid circles, this point is shared by that peg experiment.) The open circle (R_{1+2}) for $I_1 = I_2 = -0.6$ is at 42 adrians; this is then the value of the ordinate (follow the horizontal broken line). The value of the abscissa for this point was obtained as follows: the stimulus was presented in area 1 in two separate records, giving responses of 38 adrians and 40 adrians (open triangles). R_1 is taken as the average of these two responses: $R_1 = 39$ adrians. R_2 (open square) is found to be 26 adrians; hence, the abscissa is

$$R_1 + R_2 = 39 + 26 = 65 \text{ adrians.}$$

To plot the peg experiments, one must consider that two different intensities were used simultaneously. The solid triangle which is on the broken line (for which log I = -0.6) represents the response to the simultaneous presentation of stimulus 2 at -0.6 log and stimulus 1 at -2.0 log (the fixed intensity in this peg experiment); the response is 29 adrians. The separate response of area 2 alone to -0.6 log is 26 adrians (lower left, Fig. 7); the response of area 1 alone at -2.0 log is found to be the average of 17, 18, 19 and 20 adrians, or 18.5 adrians. Thus:

$$R_1 + R_2 = 26 + 18.5 = 44.5 \text{ adrians}$$

This gives the point (triangle) on the response-summation plot at (44.5, 29).

The data points shown on the response-summation plot were all derived in the fashion indicated above; the corresponding curves were similarly derived by performing the same operation on the curves fit to the data on the intensity-response plots. That the curves so generated are a good approximation to the data points is an indication that

appropriate curves were fit to the intensity-response data; the importance of this will be evident later.

The spontaneous rate is also indicated on the response-summation plot. Notice that the vertical indicator of spontaneous rate on the abscissa is at 24.2 adrians, twice the measured rate. This is because the abscissa is a sum of separate responses; each spot would "respond" to no light with the spontaneous rate, so the sum is twice the spontaneous. It is through the intersection of these indicators of the spontaneous rate (the point (24.2, 12.1) for this unit) that the line of unity slope implying linear summation must pass. Obviously, the curve in Fig. 7 is not such a line; the implications of this will be discussed later.

The Method of Sensitivity-Summation

My second method for analysing the two spot experiments is a generalization of the log I_s vs log I_d coordinates used by Easter (1968a--see "Ganglion Cells" in the Introduction) that incorporates the concept of sensitivity additivity, as used by Cleland and Enroth-Cugell (1968) and Enroth-Cugell and Pinto (1972a, b). The method consists of comparing the arithmetically summed sensitivities[g] of each area with the sensitivity of both areas stimulated simultaneously with equal intensities. If the system is linear through the final summation, the criterion of equation (3) must be met:

$$S_1 + S_2 = S_{1+2} \qquad \text{(linear)} \qquad (3)$$

where S is sensitivity. In terms of intensities:

$$\frac{1}{I_1} + \frac{1}{I_2} = \frac{1}{I_{1+2}} \qquad \text{(linear)} \qquad (4)$$

where I_1 and I_2 are the intensities in each area illuminated separately, and I_{1+2} is the intensity in each area when both are illuminated simultaneously. This condition is independent of any non-linearities after the final interaction of the two areas as separate entities. By the definition of sensitivity, the intensities in equation (4) always produced the same magnitude of response at the ganglion cell; that is, I_1 or I_2

alone elicited the same response as I_{1+2}. In Fig. 4 if R is the same, s must have been the same, regardless of the form of the function represented by the box marked F (unless it is non-monotonic).

Equation (4) should be tested for as many different levels of response as possible; to present the data, I plot them on a generalized form of Easter's (1968a) coordinates. Equation (4) may be rewritten:

$$-\log\left(\frac{1}{I_1} + \frac{1}{I_2}\right) = \log I_{1+2} \tag{5}$$

and the coordinates defined as:

$$\text{abscissa} = \log I_{1+2}$$

$$\text{ordinate} = -\log\left(\frac{1}{I_1} + \frac{1}{I_2}\right) \tag{6}$$

These I call "sensitivity-summation coordinates" and the graph a "sensitivity-summation plot." Each point on such a plot is a test of equation (5); the system is entirely linear through the final summation (at that response level) if the point lies on the line of unity slope passing through the origin (this line I indicate with a light diagonal line on each sensitivity-summation plot).

I have pointed out that the sensitivity-summation coordinates are a generalization of Easter's (1968a) coordinates. In fact, the abscissa is identical to Easter's abscissa; the ordinate makes use of sensitivity-summation to account for experiments in which the individual sensitivities of the two areas are unequal (Easter was constrained to find areas of equal sensitivity). This ordinate does not exactly simplify to Easter's ordinate for the case of equisensitive areas, for if $I_1 = I_2$:

$$\text{ordinate} = -\log\left(\frac{1}{I_s} + \frac{1}{I_s}\right) = \log I_s - \log 2$$

The difference of log 2 is a result of sensitivity-summation involving a sum, rather than a single intensity; this form is to be preferred to an average (which does reduce to Easter's coordinates) for with this system linearity is implied by the diagonal, rather than a line with an intercept of log 2.

Notice that while I refer to these coordinates as sensitivity-summation, they are in fact an intensity coordinate system. (The negative sign on the ordinate is required to restore the dimensions of intensity.) But since sensitivities on a log scale are the negatives of intensities, the coordinates may be considered either as intensity coordinates increasing in the normal fashion (to the right and up), or as sensitivity coordinates increasing to the left and downward.

Fig. 8 demonstrates the method of generating a sensitivity-summation plot from the same data as shown in Fig. 7. The plot in the upper right gives the intensity-response curves for each area stimulated separately; this is the identical plot to that in the lower left of Fig. 7. The plot on the upper left is the intensity-response plot for the equal-intensity experiment; it is identical to the plot in the upper left of Fig. 7 except that the peg experiments have been omitted.

The sensitivity-summation coordinates, defined by equations (5), appear in the lower left part of Fig. 8. The broken lines between the graphs indicate the method of locating one point on the sensitivity-summation plot, the point obtained for a response magnitude (criterion response) of 30 adrians. On the response-summation coordinates, we had found the responses for an equal intensity of stimulation (vertical broken line through the intensity-response curves); for sensitivity-summation we find the intensities required for an equal magnitude of response to the stimulation (horizontal broken lines through the intensity-response curves). Since the intensity-response data do not obligingly align themselves along lines of constant response, we must interpolate between the data points. This is accomplished by considering only the smooth curves fit to the intensity-response data; since these same curves have already been used to generate reasonable curves on the response-summation plot (Fig. 7), some faith can be placed in the propriety of the fit.

The value of $\log I_{1+2}$ is found from the intersection of the broken line at 30 adrians and the dotted curve (upper left): $\log I_{1+2} = -1.45$.

Fig. 8. Sensitivity-summation analysis of the unit shown in Fig. 7. The same stimulus is shown on the lower right. Conventions are exactly as in Fig. 7. Upper right: Intensity-response curves for each area stimulated alone; this is the same graph as on the lower left of Fig. 7. Upper left: Intensity-response curve for the equal-intensity experiment; this is the same graph as on the upper left in Fig. 7, except that the data from the peg experiments have been omitted. Lower left: Sensitivity-summation plot derived from the curves in the top graphs. Sensitivities sum linearly for any points which fall on the diagonal. The dashed lines indicate the method of locating a particular point on the sensitivity-summation plot, where $\log (I_{1+2}) = -1.45$, $-\log(\frac{1}{I_1} + \frac{1}{I_2}) = -1.26$. For details, see the text.

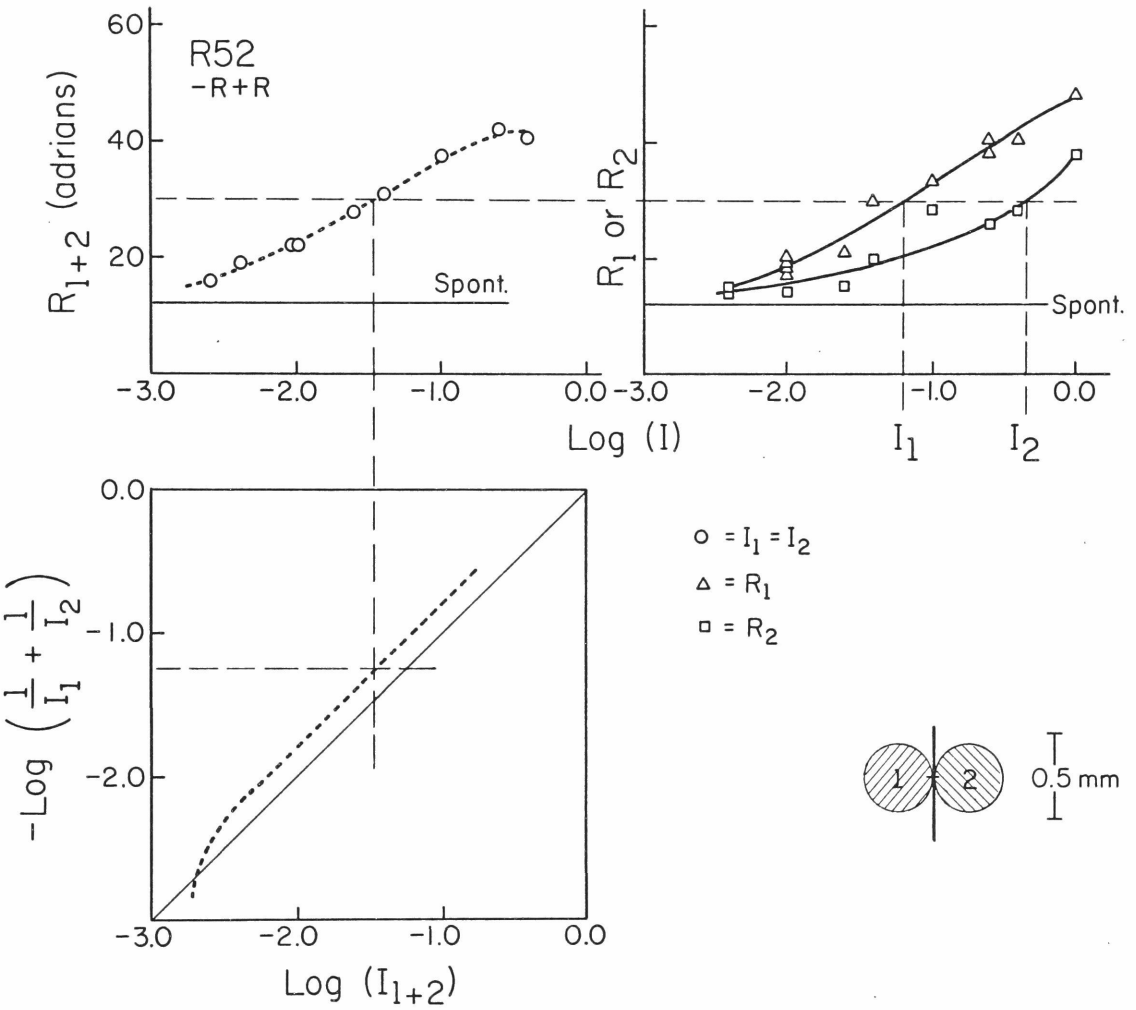
R52
-R+R
R_{1+2} (adrians)
60
40
20
Spont.
-3.0
-2.0
-1.0
0.0
R_1 or R_2
Spont.
Log (I)
I_1
I_2
$-\text{Log}\left(\frac{1}{I_1}+\frac{1}{I_2}\right)$
0.0
-1.0
-2.0
Log (I_{1+2})
o = $I_1 = I_2$
Δ = R_1
□ = R_2
1
2
0.5 mm

Fig. 8

Similarly, the intensities for the individually stimulated areas can be found (upper right):

$$\log I_1 = -1.2 \quad ; \quad \log I_2 = -0.36$$

thence:

$$\frac{1}{I_1} = 15.9 \quad ; \quad \frac{1}{I_2} = 2.3$$

and the ordinate has the value:

$$\text{ordinate} = -\log (15.9 + 2.3) = -\log(18.2) = -1.26$$

This process is repeated for as many different response levels as possible to generate the sensitivity-summation plot. The range of the plot is limited by the usable range of different responses; the highest response level for which a point can be found is the maximum firing of the less responsive area. The lowest response would be just above spontaneous, but as the intensity-response curves become extremely shallow at low response levels confidence in the exact location of the curve dwindles. As the slope of the intensity-response curve nears zero, a response difference of a fraction of an adrian can reflect an intensity difference of a significant part of a log unit.

The peg experiment is not shown on this plot. It would be simple to generalize the abscissa with an equation similar to that for the ordinate and plot the experiment, but the significance of such a curve is highly questionable. The concept of sensitivity requires finding intensities for equal responses, but it is intrinsic in the peg experiment to adhere to a particular intensity regardless of response. In such a case, the area with the lower intensity may in fact be contributing little to the response, but the reciprocal of its intensity gives it overwhelming weight.

Implications of the Response-Summation and Sensitivity-Summation Plots

The response-summation and sensitivity-summation plots were introduced as a test for non-linearities after the first interaction between the separately stimulated areas or before the final interaction of these

areas, respectively, but they contain more information than the simple linear/non-linear decision. To a considerable extent, the particular shapes of these plots can help us deduce the general form of the non-linearities that are present.

If we knew the effects on the response-summation and sensitivity-summation plots of various of the possible non-linearities postulated in the model of the center of the receptive field (Fig. 4), we might hope to identify similar features in the corresponding plots from the data. Theoretical results from hypothetical models can be found by computer simulation; I have selected particular functions and parameters for models including various non-linearities, and shall show the resultant plots drawn by computer. These computed graphs are merely illustrative; the conclusions and general principles discussed in this section are all demonstrated mathematically in Appendices II and III. In selecting the particular functional forms of non-linearities to illustrate, I have chosen what I consider the most likely possibilities (e.g.: compressive functions, inhibitory lateral interactions) but this does not exclude the possibility that we may find what I have considered the less likely alternative (e.g.: accelerated functions, lateral facilitation). The expected result from a function that operates in the opposite direction from the one hypothesized would be just the reverse of the result discussed: read "up" for "down," "larger" for "smaller," and so forth.

Before discussing particular non-linearities, we should consider the possible effect of a mismatch in sensitivity between the two areas under consideration. I shall assume that a sensitivity difference is not due to an intrinsic difference between the cones in the two areas but to a difference in the effectiveness of their contributions to the ganglion cell, such that the response from one of the areas is multiplied by a factor ρ. If there are no non-linearities after the receptors, there will be no effect of ρ on the plots; when there is a non-linearity, a non-unity ρ will always have the effect of making the system appear more linear than it is. Consider an infinite sensitivity difference.

On the response-summation plot, the response to both areas, R_{1+2}, will simply be the response to the sensitive area, which will also be the only one represented on the abscissa; a plot of R_1 vs R_1 must be a line of unity slope. Similarly, on sensitivity-summation coordinates, the sensitivity for both areas will be the sensitivity of the more sensitive, which will be the sole effective contributor to the sum on the ordinate; the plot reduces to log I_1 vs log I_1, a straight line of unity slope through the origin. (Notice that had the ordinate been defined as a simple sum of intensities ($\log(I_1+I_2)$) it would have been dominated by the much higher intensity needed in the less sensitive area, and the curve would have been displaced far above the diagonal, even if the system were linear.)

Non-linearities Before the Final Summation Three models in which there is no non-linearity after the final summation (box marked F in Fig. 4) are represented in Fig. 9. Each row of the figure shows predictions from a particular model; on the left are the intensity-response curves for each spot alone and for the equal-sensitivity experiment (peg experiment not shown), in the center are the response-summation plots showing both the equal intensity (dotted) and peg (solid) experiments, on the right are the sensitivity-summation plots for the equal-intensity experiments.

The first non-linearity we shall consider is the one assigned to the receptors (though it might occur in another unit after receptors but before any interaction). In all of the models shown in this section, I have assumed a square root function at the receptors (see "Receptors" in the Introduction). Other proposed forms of receptor non-linearities would give sensitivity-summation plots with steeper than unity slopes at high intensity levels (see Appendix III, "Alternatives to the Power Function").

The model shown in the top line of Fig. 9 is entirely linear with the exception of the square root function at the receptors. Since there is no non-linearity after the first (and only) interaction of the two

Fig. 9. Computer simulations of three hypothetical models for the center of the receptive field of a ganglion cell. In each row are the predictions for one particular model. The salient features of each model are indicated on the right; each postulated a square root function at the receptors, and no non-linearities after summation. The distinguishing features were presence or absence of thresholds and lateral interactions. The graph on the left of each row shows the intensity-response plot for each area stimulated alone (solid curves) and for the two areas stimulated simultaneously in the equal-intensity experiment (dotted). The middle graph in each row is the response-summation plot derived from the intensity-response curves; the dotted curve is for the equal-intensity experiment and the solid curve is for a peg experiment. All peg experiments were for $I_1 = -1.5$. Other details as in Fig. 7. The graph on the right in each row is the sensitivity-summation plot, conventions as in Fig. 8. Particular numerical values used in each simulation were (refer to symbol table, Appendix I):

Top row: $n = 1/2$, $\rho = 1.0$, $c = 0$, $x_o = 0$.
Middle row: $n = 1/2$, $\rho = 2.0$, $c = 0.3$, $x_o = 0$.
Bottom row: $n = 1/2$, $\rho = 2.0$, $c = 0$, $x_o = 100$.

Receptors:
$\sqrt{I}$
Post Σ:
Linear
Receptors:
$\sqrt{I}$
Thresholds=0
Post Σ:
Linear
Lateral inhibition
Receptors:
$\sqrt{I}$
Thresholds: >0
Post Σ:
Linear
— R_1, R_2
---- R_{1+2}
---- $I_1 = I_2$
— $I_1 = -1.5$
Responses (arbitrary units)
Log (I)
$(R_1 + R_2)$
(R_{1+2})
$-\text{Log}\left(\frac{I_1}{I_1} + \frac{I_1}{I_2}\right)$
Log (I_{1+2})
-3.0
-2.0
-1.0
0

Fig. 9

areas, the response-summation plots (center) are straight lines of unity slope through the origin.* (Note that in these models the spontaneous rate is assumed to be zero, so the point of zero response coincides with the origin.) The sensitivity-summation plot (right) reflects the non-linearity before the final summation; it is a line of unity slope and an intercept of 0.3, the value appropriate to a square root function (see Appendix III, "Two Spot Summation").

Now let us consider the effect of lateral interaction. To be observable at the ganglion cell, lateral interaction must be non-linear (see "A Model of Retinal Processing"); the effect of an inhibitory interaction, regardless of the particular form chosen, must be to decrease the contribution of each spot when the other is illuminated from what that spot contributes in isolation. The summation of the two areas must be less than linear; there is a decrease in the slope of the response-summation plot (not necessarily a uniform decrease), and the sensitivity-summation plot is moved downward.

As a simple demonstration, let us consider a subtractive interaction, made non-linear by the presence of thresholds (with $x_o = 0$)** in each pathway (middle row, Fig. 9: see Equations (18) and (19), pages 34 and 35). The response-summation plot is affected, as there are non-linearities (thresholds) after the first interaction; the observed effect is a decrease in slope of plots from both the equal-intensity and peg experiments. The sensitivity-summation plot is also affected, as the non-linearities are before the final summation; the curve is depressed (in this case, to slightly below the diagonal, in spite of the square root

---

*The observant reader may notice that the slope on the response-summation plot drawn by the computer is not precisely as represented. This is because the scaling of ordinate and abscissa of these plots is not quite the same.

**A threshold = 0 is a valid non-linear element, as it serves the function of blocking negative signals. A negative signal would be expected in an inhibited pathway if only the inhibiting area were illuminated.

function at the receptors). Both the change in slope on the response-summation plot and the extent of depression on the sensitivity-summation plot depend on the strength of the interaction.

Let us now remove the lateral interactions, so the two areas interact only at the main summation point (Σ in Fig. 4), and consider the effects of a non-zero threshold in each pathway. Such a model is plotted on the bottom row of Fig. 9. As in the top row, there is no non-linearity after the first interaction of the two areas, so the response-summation plots are straight lines of unity slope through the origin. The sensitivity-summation plot, however, is affected by the thresholds (a non-linearity before the final summation); the effect is most pronounced at low levels when the cone outputs are of the same general magnitude as the threshold level, x_o. The manifestation is a depression of the sensitivity-summation plot at the lowest intensities (see "Thresholds," Appendix III). The depression of the sensitivity-summation plot due to thresholds would be confounded with that due to inhibitory lateral interactions, were it not for the fact that thresholds have no effect on the response-summation plot while lateral interactions decrease its slope.

Non-linearities After the Final Summation Let us now turn our attention to the effects of a non-linear function after summation, the box F in Fig. 4. Such a function cannot have an effect on the sensitivity-summation plot for it is after the final interaction of the two areas, but only on the response-summation plots.

The curves in the top row of Fig. 10 were derived from a model in which the transform F had the form of a power function:

$$R = g \cdot s^{\gamma}$$

where g and γ are constants. The effect upon the response-summation plot is a change in slope of the line for the equal-intensity experiment. A power function is the only function for which a straight line results; the linear function after summation is a special case of a power function, in which $\gamma = 1$. The peg experiment, however, does not plot

Fig. 10. Computer simulations of three more hypothetical models. Again, each model assumed a square root function at the receptors: the distinguishing features were the non-linear function after summation, and the thresholds. Conventions and organization are the same as in Fig. 9. (Peg experiments are again at $I_1 = -1.5$, with the exception of the bottom row, for which the peg experiment was for $I_1 = -2.6$.) Numerical values used were (see symbol table, Appendix I):

Top row: $n = 1/2$, $\rho = 0.75$, $c = 0$, $x_o = 0$, $\gamma = 0.5$.
Middle row: $n = 1/2$, $\rho = 2.0$, $c = 0$, $x_o = 0$, $k = 750$.
Bottom row: $n = 1/2$, $\rho = 2.0$, $c = 0$, $x_o = 130$, $k = 1000$.

Receptors:
$\sqrt{I}$
Post Σ:
Non-linear
(power)
Receptors:
$\sqrt{I}$
Post Σ:
Non-linear
Receptors:
$\sqrt{I}$
Thresholds: >0
Post Σ:
Non-linear
R_1, R_2
R_{1+2}
$I_1 = I_2$
$I_1 = -1.5$
$I = -2.6$
$-\log\left(\frac{1}{I_1^1} + \frac{1}{I_2^1}\right)$
$(R_1 + R_2)$
(R_{1+2})
Log (I)
Log (I_{1+2})
Responses (arbitrary units)
-3.0
-2.0
-1.0
0

Fig. 10

as a straight line if $\gamma \neq 1$; for the particular power used in this example (1/2) the curve is an hyperbola tangent to the equal-intensity curve at the fixed value of I_1.

The middle row of Fig. 10 shows the effect of another non-linear function, which I call a compressive function[g]. A compressive function is one that is nearly linear at low values, but breaks away from linearity at higher values, the output ultimately reaching an asymptote which it can never exceed. The particular simple compressive function used in these simulations has the form:

$$R = \frac{g \cdot s}{k + s} \qquad (7)$$

where g and k are constants. The response-summation plot from the equal-intensity experiment is a curve, which can be shown to be independent of the value of k. At low responses, the curve approximates a straight line of unity slope through the origin; at higher levels it breaks away to shallower slopes: this curve itself is described by an equation of the same general form as (7) (see Appendix II). The peg experiment gives a rather different curve. Notice also that this model is the first one discussed for which the intensity-response curves (left column, middle row) have the s-shape normally observed in the data. The power function at the receptors results in a curve that is concave upwards on the intensity-response plot; the compressive function causes the curve to become concave downward at the highest levels.

Finally, the bottom line of Fig. 10 demonstrates a model combining the compressive function of the immediately preceding model with the thresholds of the last model in Fig. 9. Each of these non-linearities expresses itself in its characteristic way: the thresholds depress the sensitivity-summation plot at lowest intensities, while at higher levels it tends toward a line whose intercept is determined by the receptor function; the compressive function results in a curved response-summation plot. This response-summation plot is not different from the previous plot despite the choice of a different value of k; it merely covers a larger range of the same curve (see Appendix II). A peg experiment is

also shown, in this case for a low fixed intensity. The similarity of the predictions of this model and the data shown in Fig. 7 and Fig. 8 should be noted; the same general features will appear in data shown throughout the remainder of this thesis.

Summary of Response-Summation and Sensitivity-Summation

General features of the response-summation plot:

1) It is unaffected by non-linearities before the first interaction between the two areas.

2) It indicates linearity of all stages after the first interaction by a straight line of unity slope through the zero-response point.

3) It is affected only by non-linearities at or after the first interaction of the two areas.

4) It indicates inhibitory lateral interactions by a decrease in slope.

5) It indicates non-linearity after summation by a change in slope if the non-linearity is a power function, by a curve for other non-linearities.

* * *

General features of the sensitivity-summation plot:

1) It is unaffected by non-linearities after the final interaction of the two areas.

2) It indicates linearity of all stages prior to the final summation by a straight line of unity slope through the origin.

3) It is affected only by non-linearities at or before the final interaction of the two areas.

4) It indicates inhibitory lateral interaction by a depression of the curve, with or without a slope change.

5) It indicates a power function before the first interaction of separate areas by a straight line of unity slope. The intercept of the line is related to the exponent of the power function. Non-linearities other than power functions do not give straight lines of unity slope.

* * *

The potential confusion among possibilities presented by the sensitivity-summation plot may often be resolved by comparison with the response-summation plot. For instance, one could decide whether or not lateral interactions were affecting the intercept by observing the slope of the response-summation plot.

EXPERIMENTAL RESULTS AND CONCLUSIONS

We are now prepared to examine the results of two spot experiments performed on either the centers or surrounds of the receptive fields of goldfish ganglion cells. In this chapter I shall present the results from a number of such experiments; following the results, I shall indicate the conclusions which may be drawn from them. In general, the particular set of data shown will illustrate the typical outcome of the experiment under consideration; the effect will have been replicated in other units (see Table I). In some cases, the same experiment on other units may have shown a less pronounced effect, or even failed to show any effect (in the latter case, the reasons for this lapse will be discussed); in no case were there ever contradictory results.

I shall first show the general results obtained, and discuss their relationship to the model shown in Fig. 4. Next I shall concentrate on confirming the conclusions drawn about interactions in the retina; finally, I shall extend the analysis to temporal relationships in the retina. Before beginning, however, let me point out it is essential that the reader have a thorough grasp of the response-summation and sensitivity-summation methods discussed in the previous chapter.

General Results from the Two Spot Experiments

I have selected experiments on six units to illustrate the general results of the two spot experiments. The experiments were chosen to encompass as many different stimulus conditions and response types as possible. The reader is encouraged also to refer back to the unit shown in Figs. 7 and 8, and to notice these same features in the response-summation and sensitivity-summation plots presented in subsequent sections (Figs. 14, 15, 16, 17, 19).

Response-Summation Plots Response-summation plots for the six representative units are shown in Fig. 11 (A-F). Each unit's response type and a plan of the stimulus configuration used is shown in the upper left corner of its plot. The scaling of the axes is identical in all six plots.

Several different stimuli were used in these six experiments. The most common was two small spots both within the center of the receptive field; the spots shown were 0.2 mm in diameter (Fig. 11, A and D) and 0.4 mm in diameter (Fig. 11, B and C). Typically, the spots were placed symmetrically about the middle of the receptive field; an exception is Fig. 11 C. In Fig. 11 E the stimulus used was an 0.4 mm diameter spot and a concentric annulus of the same area; both spot and annulus lay entirely within the center of the receptive field. Only one unit is shown in this figure for which the stimuli were not within the center: the two partial annuli in Fig. 11 F each lay entirely within the surround of the receptive field of that unit. (See also Fig. 14.)

The spontaneous firing rates of the six units are indicated by the solid horizontal and vertical lines (see "Method of Response-Summation," above). With a single exception, which lacked any spontaneous firing (Fig. 11 C), the spontaneous rates of these units were about 13 adrians, ranging from 11 to 14 adrians. The range of response rates represented is large. For the stimuli used, the maximal response from most of the units was about 60 adrians, with a range of 80 adrians (Fig. 11 F) to 20 adrians (Fig. 11 C). Most of the responses were "off" responses, the exceptions being the center "on" response in Fig. 11 B, and the surround "on" response in Fig. 11 F.

There was no sensitivity difference between the two areas stimulated for the unit of Fig. 11 A, but in most cases the sensitivities of the two areas were not the same. The two areas in Fig. 11 C showed a sensitivity difference of about 0.7 log units; that is, the light in one area had to be made 0.7 log units more intense than the light in the other area to achieve the same response. Similarly, the areas in Fig. 11 E differed by 0.6 log units; the other units shown in the figure were in the more usual range of 0.2 to 0.3 log units sensitivity difference.

In view of the broad variety of stimuli and responses represented by these six units, it is remarkable how similar are the features of the

Fig. 11. Response-summation plots from six representative units. Conventions as in Fig. 7. All stimuli were 710 nm.

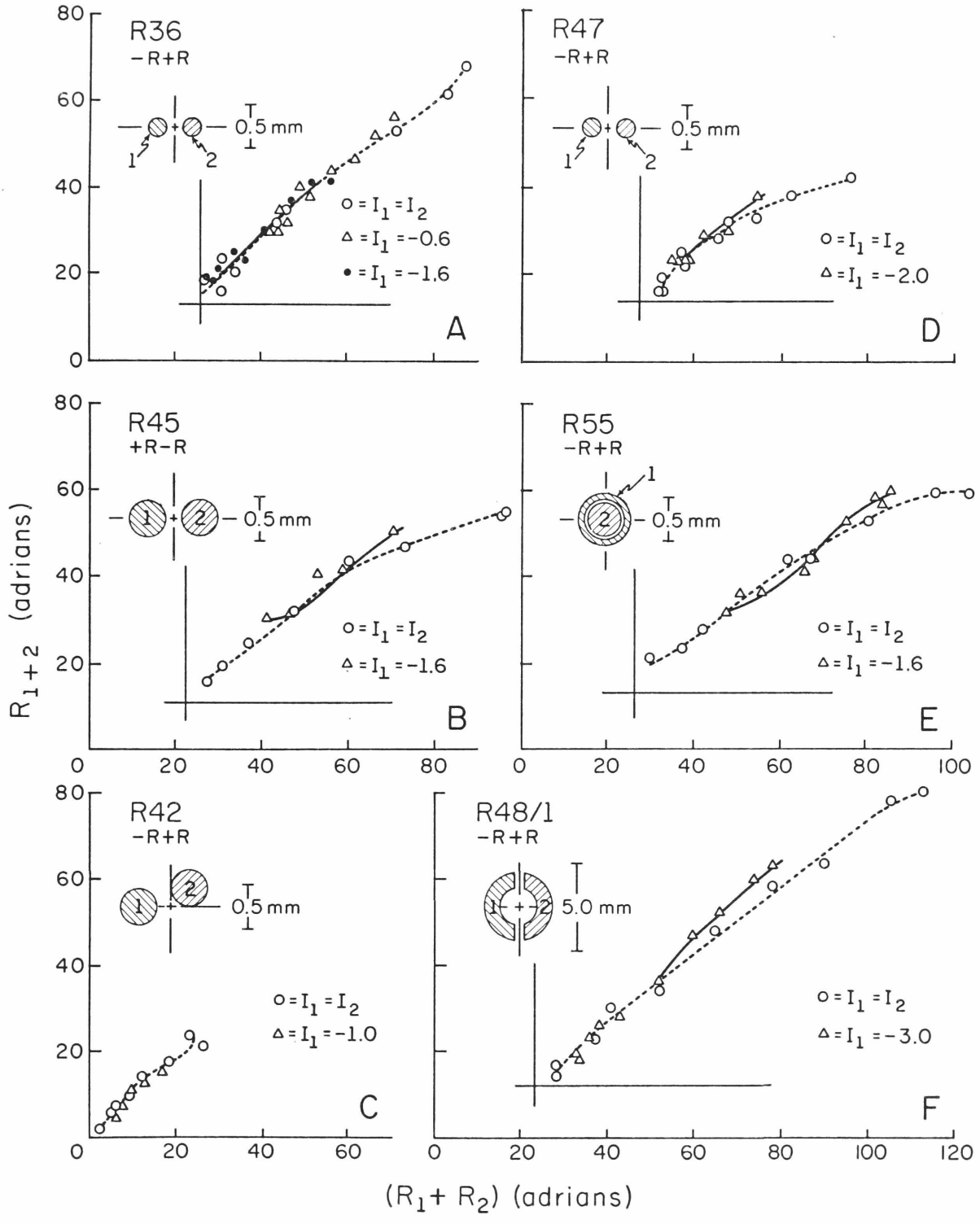
R36
−R+R
0.5 mm
o = I1 = I2
Δ = I1 = −0.6
• = I1 = −1.6
A
R47
−R+R
0.5 mm
o = I1 = I2
Δ = I1 = −2.0
D
R45
+R−R
0.5 mm
o = I1 = I2
Δ = I1 = −1.6
B
R55
−R+R
0.5 mm
o = I1 = I2
Δ = I1 = −1.6
E
R42
−R+R
0.5 mm
o = I1 = I2
Δ = I1 = −1.0
C
R48/1
−R+R
5.0 mm
o = I1 = I2
Δ = I1 = −3.0
F
R1+2 (adrians)
(R1 + R2) (adrians)

Fig. 11

six response-summation plots. The curves for the equal-intensity experiments (dotted curves in Fig. 11) all approximate a line of unity slope through the zero response point in the vicinity of that point. At higher firing levels, the curves tend to depart from this line, approaching a line of shallower slope. In some cases, the departure is quite pronounced (Fig. 11 B, D, E); in some it is less pronounced (Fig. 11 A, C, F).

A number of peg experiments are also shown in Fig. 11 (triangles or solid circles, and solid curves). Peg experiments in which the fixed intensity was sufficiently low that the points fell within the linear portion of the response-summation plot tended to coincide with the curves from the equal-intensity experiment (Fig. A (solid circles), C); peg experiments at higher intensities tend not to coincide with the curves from the equal-intensity experiment (Fig. 11 B, D, E, F).

Since the response-summation plots depart from a straight line of unity slope through the point of zero response, there must be a non-linearity after the first level of interaction between the two areas. That the departure from linearity occurs only at higher response levels implies that the non-linearity is of the compressive[g] type. It is not a power function, or the curve would still be a straight line, albeit of less than unity slope (see "Non-linearities After Final Summation" in Methods of Analysis); it must be a function that is linear for low levels, and becomes progressively more compressed at higher levels.

The results of the peg experiments also tend to confirm this hypothesis. A non-linearity after summation has the effect of causing the curves from the peg experiments to depart from the curves due to the equal-intensity experiments. A non-linear interaction prior to summation, which might have been invoked to explain the departure from linearity, need not cause this difference between curves from the equal-intensity and peg experiments. The reader is invited to compare these results with the response-summation plots predicted for a model with a non-linearity after the summation, Fig. 10.

Sensitivity-Summation Plots The sensitivity-summation plots for the six units of Fig. 11 are shown in Fig. 12. These plots are derived from the same intensity-response curves as were used to generate the dotted curves of Fig. 11, so the same diversity of stimuli and responses underlies these curves. Despite these differences, the salient features of the sensitivity-summation plots are quite similar. At higher intensities, each approaches a straight segment of unity slope with an intercept between 0.15 and 0.3 log units. With the exception of Fig. 12 A, each curve is depressed at the lower intensity levels.

A straight line of unity slope on sensitivity-summation coordinates implies there is a function which may be described by a power relationship before the final summation. The value of the exponent of this power function may be derived from the line's intercept with the ordinate, but corrections must be made for sensitivity differences between the two areas stimulated in the experiment. I have made estimates of the presumed exponent using an iterative procedure which corrects the form of the ordinate (see "Estimation of the Receptor Exponent, n" in Appendix III). Five iterations were applied to each of 19 separate experiments on 14 different units, including the six shown in Fig. 12. (Often after the completion of an experiment a new pair of stimuli was set up, and another experiment performed in which different sub-areas of the receptive field were stimulated.) The values of the exponent ranged from 0.41 to 0.69, with a mean of 0.54. This value is sufficiently close to 1/2 that I shall speak of this as a square root function. Since one expects a non-linearity at the receptors, I shall assign this function to them; however, the experiment demonstrates merely that there is such a function somewhere before the summation of the sub-areas.

The depression of the sensitivity-response plot at low intensities requires additional explanation. There are two general classes of the curve at low intensities: in some cells, it tends asympotically to the diagonal (Fig. 12 B, E), while in others it actually crosses the diagonal (Fig. 12 C, F). One or the other of these tendencies is evidenced

Fig. 12 . Sensitivity-summation plots from the same six representative units as Fig. 11. Conventions as in Fig. 8. (Note that these plots were derived from the same intensity-response curves as were used to generate the continuous curves in Fig. 11.)

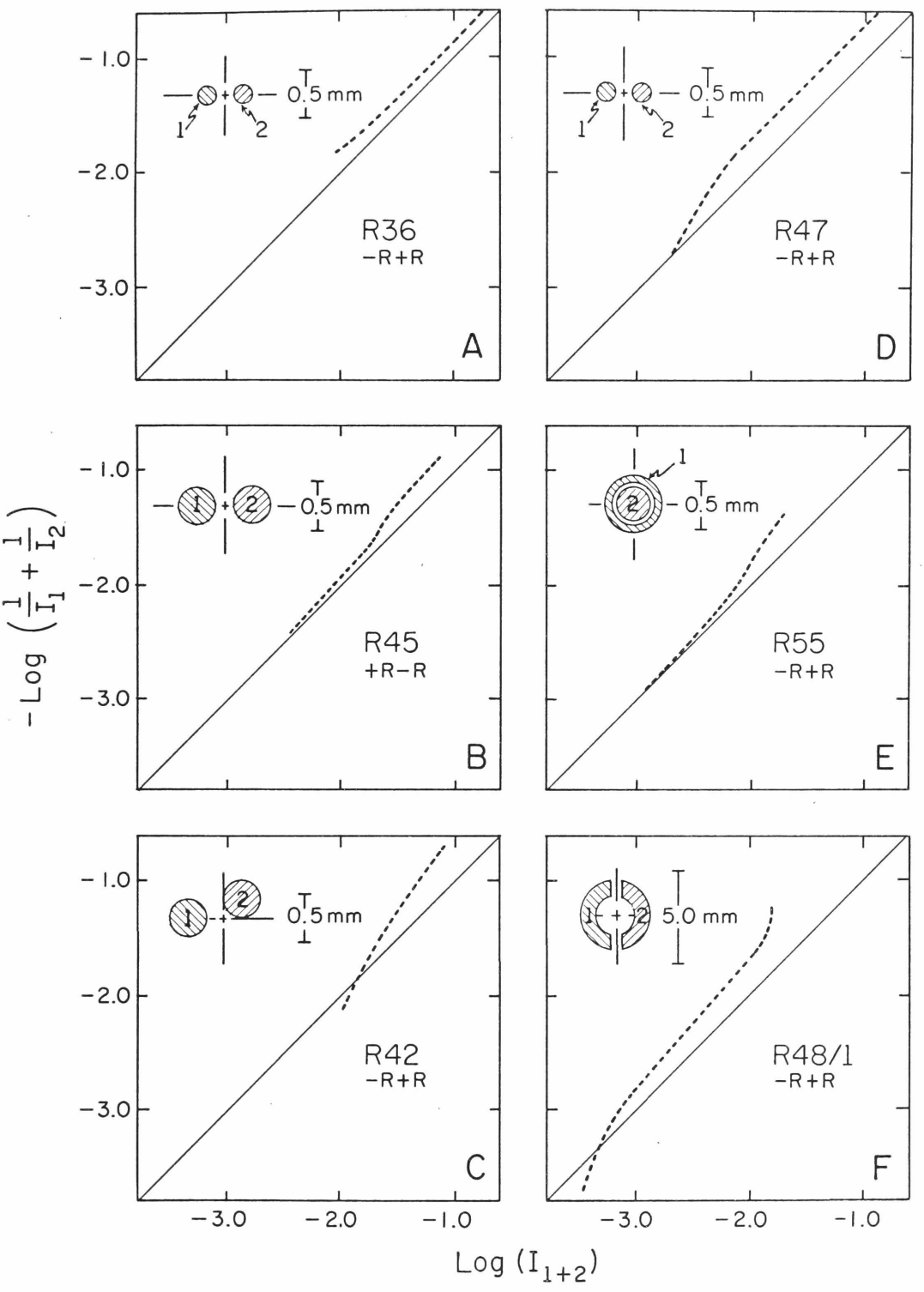
-1.0
-2.0
-3.0
0.5 mm
1
2
R36
-R+R
A
R47
-R+R
D
R45
+R-R
B
R55
-R+R
E
R42
-R+R
C
5.0 mm
R48/1
-R+R
F
-3.0
-2.0
-1.0
$-\text{Log}\left(\frac{1}{I_1}+\frac{1}{I_2}\right)$
$\text{Log}\,(I_{1+2})$

Fig. 12

in most of the other units not included in this figure. The curve crossing the diagonal can be explained by postulating a threshold in the pathway from each group of receptors before summation. (See "Thresholds" in Appendix III; also refer to Fig. 10, bottom row, and Fig. 4.) An asymptote with the diagonal can be explained by assuming that the power function at the receptors fails at low intensities, becoming more linear. A plausible model of the receptor function would be one in which the square root function is generated by having the output of the receptor divide the input to it; an electrical analog would be a shunting[g] feedback with a small constant "leakage," ℓ. For an input intensity I, the receptor output, x, would be:

$$x = \frac{I}{x + \ell} \qquad (1)$$

If x is large compared to ℓ, this becomes:

$$x = \frac{I}{x}$$

or

$$x = \sqrt{I} \ .$$

But if x is small compared to ℓ, (at low intensities) equation (1) becomes:

$$x = I/\ell$$

which is linear. With no non-linearity before the final summation, the points on the sensitivity-summation plot must fall along the diagonal.

The two possibilities described above are not mutually exclusive. It is quite reasonable to assume that the receptor function becomes linear at low intensities, and also that there are thresholds which must be exceeded before the contribution from a given area will be included in the ganglion cell's response. Which effect determines the form of the sensitivity-summation plot in any given experiment would then depend on the relationship between the thresholds and the level at which the receptor function becomes linear. Lest too much be made of this, the reader is reminded that sensitivity-summation plots are relatively unreliable at low light levels, owing to the extreme shallowness of the intensity-response curves (see "Method of Sensitivity-Summation").

One other feature occasionally observed is a tendency for the sensitivity-summation curves to turn upward at the highest intensities (Fig. 12 F). Like the tendency for the curve to asymptote to the diagonal at low intensities this implies a deviation from the square root function at the receptors, in the direction of still greater compression at the highest intensities. The square root function must fail at sufficiently high intensities, for the receptor cannot achieve infinite output levels. That this feature is not generally observed may be explained as follows: assume the compressive function after summation usually limits the ganglion cell output to levels lower than that for which the receptors depart from the square root function. Those few units in which the receptor asymptote is observed may be units for which the compression after summation takes effect at relatively higher levels.

Data from 24 separate two spot experiments on 17 units are summarized in Table I. The units' designation, its response type, and the particular stimulus used are indicated in the three leftmost columns. If several units were recorded from the same retina, they are distinguished as "/1," "/2;" in the cases of R43 and R48 the various units were all recorded in the same location, and may have in fact been the same cell. These cells were designated as separate units because their sensitivities were found to be different from that originally recorded. (Sudden changes in a unit's responses during the course of an experiment were also noted by Spekreijse et al., 1972.)

The middle two columns summarize the results on the response-summation plots. The first of these columns (POST FN) reports the presumed nature of the function after summation (F, in Fig. 4); a power function (P) is indicated if the equal-intensity experiment plots as an approximately straight line of less than unity slope, while a compressive function (C) is indicated if the plot is a curve. The second column (PEG) describes the behavior of the peg experiments, with regard to whether or not the plots for these experiments superimposed on or deviated from the curve for the equal-intensity experiment.

Table I

Summary of Results on Response-Summation and Sensitivity-Summation Plots

CELL*	TYPE*	STIM≠	RESPONSE SUMMATION		SENSITIVITY SUMMATION		
			POST FN†	PEG+	REC EXP∡	REC SAT∅	LOW LEVEL∉
R36	-R+R	cc1	P? C?	LN,HN	0.69	Y	0
R39	-R+R	cc1	P?	LN	0.54	Y	0
R42	-R+R	ss	C	HY	0.50	N	3?
R42	-R+R	cc1	C	LN	0.49	N	3
R43/1	+R-R	cc1	C	LY	?	N	3
R43/2	+R-R	cc1	C	LN	0.43	N	3
R43/3	+R-R	cc1	C? P?	LN	0.43	N	3
R44	-R+R	cc1	C	LY?	0.60	N	2
R45	+R-R	cc1	C	LY	0.59	N	2?
R47	-R+R	cc1	C	LY	0.41	N	1
R48/1	-R+R	ss	P? C?	LY	0.51	Y	3
R48/2	-R+R	ss	C	LY,HY	0.62	Y?	0
R52	-R+R	cc1	C	LN,HY	0.55	N	1
R53/1	+R-R	cc1	C	LN,HY	0.57	Y	3
R53/2	-R+R	cc1N	P	LN,HY	?	N	3
R53/2	-R+R	cc1F	C	-	?	?	?
R54	-R+R	cc2S	P? C?	LN	0.57	N	3
R54	-R+R	cc1F	C? P?	LN	0.55	N	3?
R54	-R+R	cc1N	C?	-	0.67	N	3?
R54	-R+R	cc2L	C	-	0.43	N	0
R55	-R+R	cc2	C	LN?	0.43	N	2
R55	-R+R	cc3N	C	LN	0.56	N	0
R55	-R+R	cc3F	C	LN	0.56	N	3
R56	-R+R	cc2	P?	-	062	N	0

Notes: *Cell identification and type as given in Fig. 7.

≠Stimulus configurations: cc1=two spots, both in center of receptive field; cc2=spot and contiguous annulus, both in center of receptive field; cc3=cross globicus; ss=half annuli, each within the surround of the receptive field. N,F,S, and L indicate spots near each other or well separated, and small or large stimuli, respectively.

†P = power function or log function after summation, as indicated by straight line on response-summation plot. C = compressive function after summation.

+LN = low peg value whose curve does not deviate from the curve for the equal-intensity experiment. LY = low peg value whose curve does depart from the curve for equal-intensity. HN = high peg value whose curve does not deviate; HY = high peg whose curve does deviate.

∡ Exponents for receptor function found by the iterative method described in Appendix III ("Estimation of Receptor Exponent, n")

∅ Y = sensitivity-summation plot shows upturn at high intensity end, indicating receptor saturation. N = no upturn evident.

∉ Indicates behavior of sensitivity-summation plot for low intensities. 0 = sensitivity-summation plot remains parallel to the diagonal at low intensities. 1 = sensitivity-summation plot depressed at low intensities, but not clear whether or not curve would cross the diagonal. 2 = curve coincides with diagonal at low intensities. 3 = curve crosses diagonal.

The three columns on the right summarize the results on the sensitivity-summation plots. The first of these (REC EXP) is the estimated exponent (n) of the presumed power function at the receptors (E, in Fig. 4). The estimates were based on the intercept of the straight line portion of the sensitivity-summation plot, using the iterative method derived in Appendix III ("Estimation of Receptor Exponent, n"). The next column (REC SAT) reports whether or not an upturn was noted at the high intensity end of the sensitivity-summation plot. The final column (LOW LEVEL) describes the form of the sensitivity-summation plot at low intensities.

Summary of the Receptive Field Center We may now apply the hypothesized functions assumed from the response-summation and sensitivity-summation plots to the model of the center of the receptive field (Fig. 4). This is also a model of the surround of the receptive field, for the same results were obtained when both stimulated areas were in the periphery, Fig. 11 F, Fig. 12 F. The revised form of the model is shown in Fig. 13.

The light incident on each group of receptors is transformed into a signal, x, where x is approximately the square root of the intensity. This relationship is sketched on the right in Fig. 13, showing x as a function of log I; the dashed portion of the curve indicates that the receptor relationship probably departs from a square root relationship at high intensities. Neither this compression nor the possibility that the receptor function becomes linear at low intensities is indicated by the first equation on the left in Fig. 13.

To be included in the sum, the output of a group of receptors must exceed some threshold value, x_o; a threshold is indicated in each pathway before summation. The input-output relationships of the thresholds are sketched to the right, and defined by the equations to the left. Notice that in the absence of any negative influences (e.g. lateral interactions) the thresholds would have no effect unless $x_o > 0$, as it is shown in the figure.

Fig. 13. Schematic of the model found to describe satisfactorily the summation between sub-areas within either the center or the surround of the receptive field (long-wavelength cones only). The mathematical functions presumed to apply at each level are indicated on the left; graphic sketches of these functions appear on the right. Conventions and symbols as in Fig. 4.

$x \approx \sqrt{I}$

$y = \begin{cases} x - x_0, & x > x_0 \\ 0, & x \leq x_0 \end{cases}$

$S = y_1 + y_2$

$R \approx g \frac{S}{K + S}$

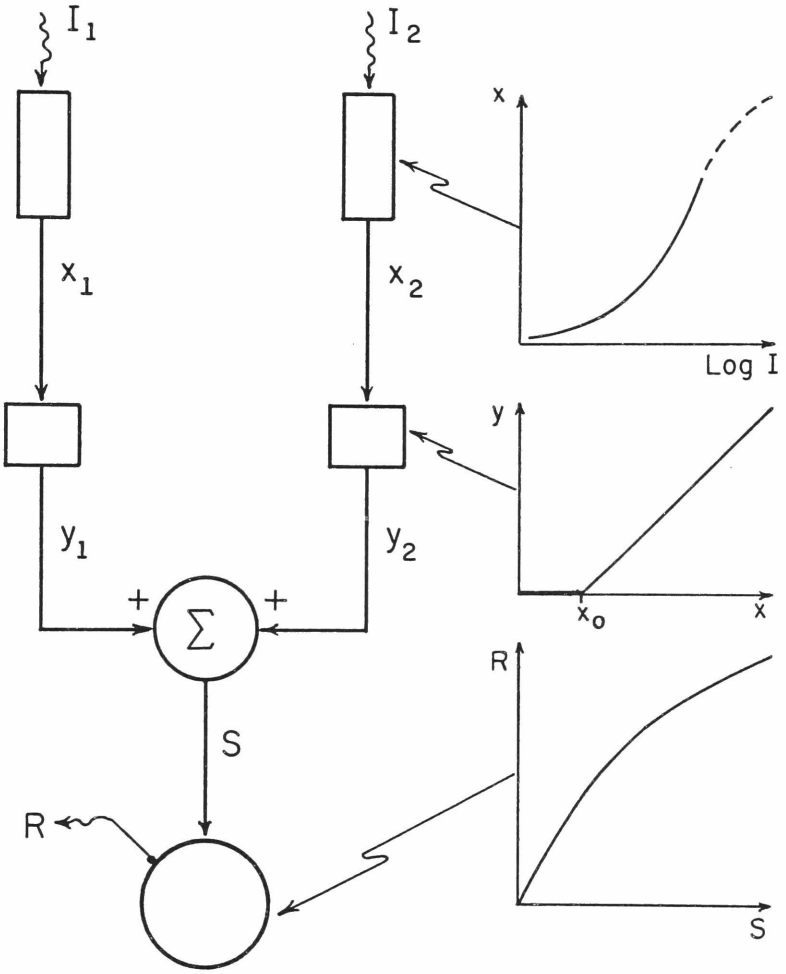

Fig. 13

The approximately square root function at the receptors, combined with the thresholds, is sufficient to explain the form of the sensitivity-summation plots in Fig. 12. The square root function accounts for a straight line of unity slope with an intercept of about 0.3 log units on the sensitivity-summation coordinates; the saturation of the receptors at high intensity accounts for the slight upturn occasionally observed at the highest response criteria. Linearization of the receptor function at low intensities may account for the plot approaching the diagonal at lower intensities, but thresholds are necessary to account for the sensitivity-summation plot crossing the diagonal.

I have assumed that the influences of the two areas then sum linearly (see "Model of Retinal Processing" in the Introduction); following summation, there is another compressive function (Fig. 13). Such a compressive function as sketched on the right might be described by an equation of the form on the left. This late compression accounts for the non-linearity observed on the response-summation plots, Fig. 11.

The model in Fig. 13 is thus sufficient to describe summation of separate areas within either the center or the periphery of the receptive field of goldfish ganglion cells (when only the long-wavelength sensitive cones are stimulated). One striking feature of the model is the complete lack of lateral interactions; the only interaction between receptors in different areas is the simple summation of their influences at the summation point. This aspect of the model I shall examine closely in the next section.

Lateral Interactions

In discussing lateral interactions, we must bear in mind that an entirely linear interaction would be undetectable at the ganglion cell, save perhaps by exploiting some difference between the time courses of interactions and summations. (For a discussion of this point, refer to "A Model of Retinal Processing," in the Introduction; equation (16) describes this situation.) Linear interactions might in fact be present, and might be recorded from the appropriate distal units, but they are

apparently not necessary for this minimal model of the processing leading to the ganglion cell responses. This section is concerned with the possibility that a non-linear interaction significant to the ganglion cell has been overlooked in the previous analysis. I shall approach this possibility from three independent points; as we shall see, each of these analyses leads to the same conclusion: there is no need to postulate lateral interactions between two separate areas.

Interaction vs Compression After Summation The first possibility to consider is that the non-linearity giving rise to the curvature of the functions on the response-summation plots is not caused by a compression after summation, but by a non-linear interaction before summation. We have noted (Fig. 9) that one specific form of interaction, a linear interaction made non-linear by the inclusion of thresholds, results in a decrease of slope in the response-summation plot--but for this form there is no change in the curvature. Might not other forms of non-linear interaction, such as a multiplicative interaction, give rise to the curvature we have observed in the response-summation plots? One form of "multiplicative forward inhibition" that has been postulated (Furman, 1965; Grüsser et al., 1970) has the form:

$$y_j = \frac{x_j}{k + x_i + x_j + x_k + \ldots}$$

where the x's are the outputs of all active elements (receptors) and k is a constant. That is, the response from each area is divided by the responses from all active areas, including itself. But this is exactly identical to a compressive function after summation, for:

$$R = \frac{(x_1 + x_2)}{k + (x_1 + x_2)} = \frac{s}{k + s} \qquad (2)$$

Since this is mathematically equivalent, it must be indistinguishable from compression after summation; in fact, this could be the mechanism by which the compression is achieved.

Let us now consider an interaction which is not equivalent to compression after summation; an example might be a divisive interaction in which the area itself is not included in the denominator. Such an interaction could also account for a deviation from linearity on the response-summation plot. This non-linearity occurs before the final summation, and is not mathematically equivalent to a non-linearity wholly after the summation; hence it must also affect the sensitivity-summation plot. Let us see if there is indeed such an effect, or if one might have been obscured by the irregularities of the sensitivity-summation plot caused by changes in the receptor function or the influence of thresholds.

Fig. 14 demonstrates two attempts to find an effect in the sensitivity-summation plot corresponding to the curvature of the line on the response-summation plot. Two different units are illustrated; the unit on the left responded at "off" to stimulation in the center of the receptive field while the one on the right gave "on" responses to stimulation in the periphery.

Each of the response-summation plots on the bottom of Fig. 14 shows a fairly sharp break away from unity slope at a point indicated by the vertical arrow. (Recall that the dotted curves were derived from the curves fit to the data points on the response-summation plots.) This break could be due to one of two causes: 1) either the total response (physiological sum) is compressed when responses exceed some level; or, 2) one (or both) of the individual intensity-response curves (R_1 or R_2) is accelerated more rapidly above some response level, without a concomitant acceleration of the curve for both together. The latter possibility is less likely, but even if it were the case, there would have to be a compression of the response to both in order to compensate for the acceleration of the receptor functions. It is therefore meaningful to identify the point at which the response-summation curve breaks from linearity with a response level (R_{1+2}); the response level at which this occurs can be located on the intensity-response curve for the equal-intensity experiment shown at the top of Fig. 14. This point

Fig. 14. Comparison, for two cells, of response-summation and sensitivity-summation plots with special reference to the point at which the curve on the response-summation plot breaks from linearity (deviates from unity slope). Stimuli (710 nm) are sketched at the top. Responses to each area stimulated alone are indicated by solid circles and triangles; all other conventions and symbols are the same as in Figs. 7 and 8. The top graph in each column shows the intensity-response curves for each area alone (solid symbols) and for the equal-intensity experiment (open circles). The middle graph in each is the sensitivity-summation plot; the bottom graph is the response-summation plot. The continuous functions on the response-summation and sensitivity-summation plots were derived from the curves fit to the intensity-response data for each unit. The vertical arrows indicate corresponding points on the three forms of plot (see text).

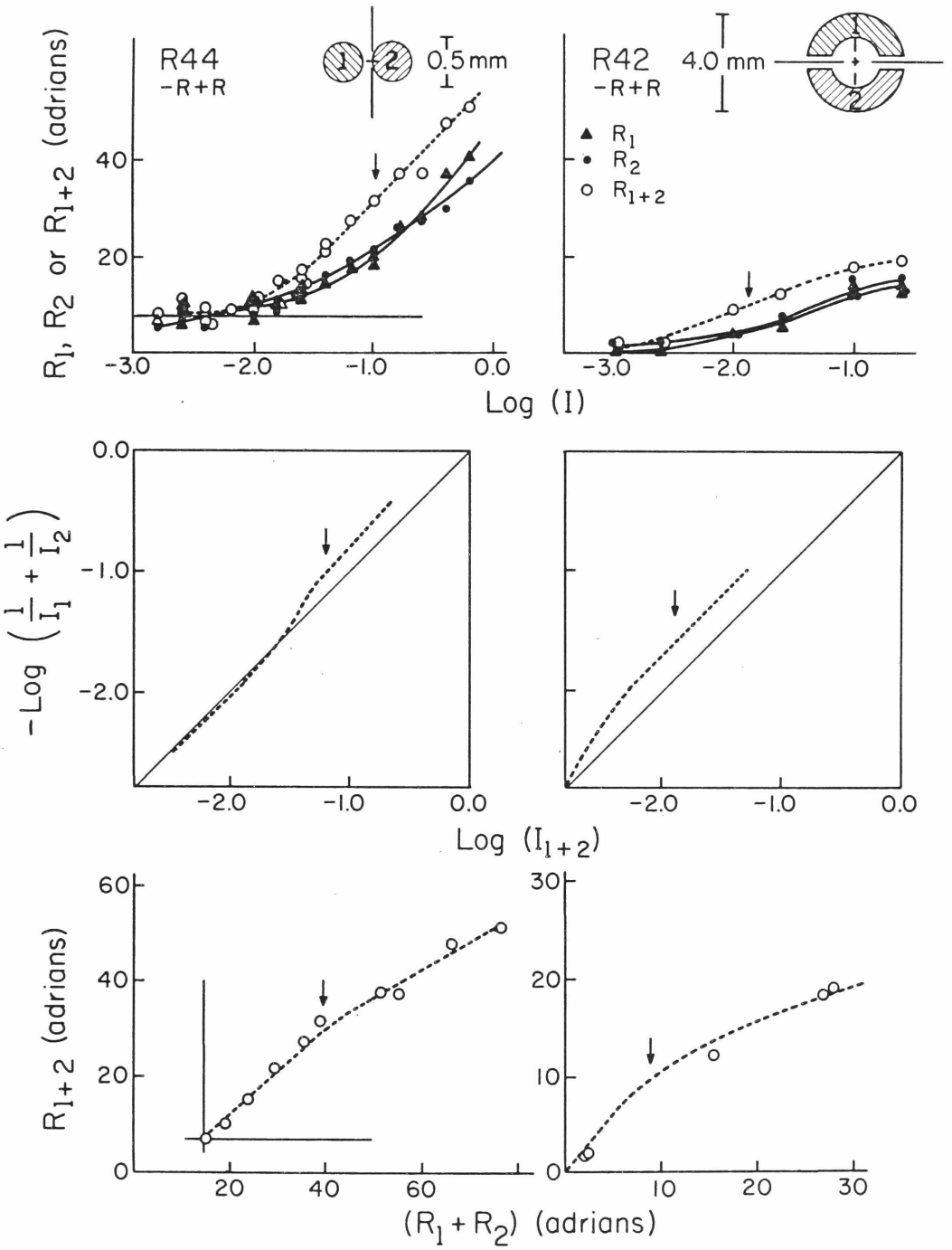
R44
-R+R
0.5 mm
R42
-R+R
4.0 mm
R1
R2
R1+2
R1, R2 or R1+2 (adrians)
Log (I)
-Log (1/I1 + 1/I2)
Log (I1+2)
R1+2 (adrians)
(R1 + R2) (adrians)

Fig. 14

(arrow) identifies the intensity (I_{1+2}) at which the break occurs. Since the abscissa of the sensitivity-summation plot is also log I_{1+2}, it is possible to locate the corresponding point on the sensitivity-summation plot (middle of Fig. 14; arrow).

We have thus located the point on the sensitivity-summation plot which corresponds to the point at which the response-summation plot breaks from linearity. In both units shown in Fig. 14, this point is well within the segment of the curve which is still a straight line. The effect of non-linear lateral interactions on the sensitivity-summation plot is to shift the curve; at the point corresponding to the break from linearity on the response-summation plot, there is no change in the sensitivity-summation plot. Therefore, if the curvature on the response-summation plot is due to an interaction prior to summation, the interaction must be equivalent to a compressive function after summation. The model as shown in Fig. 13 requires no modification.

The analysis shown in Fig. 14 could be performed on only a few units. In most units, the corresponding point was at a higher value of I_{1+2} than the maximum for which the sensitivity-summation plot could be drawn; the sensitivity-summation plot is limited, since the highest response criterion which can be used is the maximum firing of the least responsive area. However, a number of other units tended toward the same outcome, and no unit's responses contradicted the above conclusions.

Distance Dependence Let us now entertain the possibility that in addition to the non-linearities shown in the model of Fig. 13, there are non-linear lateral interactions which have escaped our notice (see "A Model of Retinal Processing"). On the sensitivity-summation plot, inhibitory interactions would depress the curve. On the response-summation plot, inhibitory interactions would decrease the slope of the curve, but the decrease might be unnoticeable in the short region before the compressive function after summation expresses itself. (Note that facilitatory interactions would have just the opposite effects.)

We shall take advantage of one of the expected properties of lateral interactions in order to try to demonstrate their presence: magnitude of interaction should change with the separation between the interacting areas. We shall compare results from stimuli that are close together with those from stimuli that are well separated, and see if there is a difference between the plots from each case.

An experiment designed to maximize and then minimize lateral interactions is shown in Fig. 15. The stimuli are shown at the upper right; each is a four-spoked pinwheel correctly termed a cross globicus (Franklyn, 1965) centered on, and confined to, the center of the receptive field. One cross, shown lightly stippled, was held in one position throughout the experiment. The other cross, shaded with lines in the diagram, could take either of two positions; in one position it was rotated until it was as close as possible to the stationary cross without overlapping it, while in the other position it was as well separated as possible. The crosses take advantage of the approximately radial symmetry of the receptive field; they may be rotated without significantly changing their sensitivities. They also provide a relatively large boundary along which interactions may occur when the two crosses are nearly adjacent.

The intensity-response functions for each cross presented alone are shown in the lower left of Fig. 15. Open circles represent the stationary cross, solid circles represent the moveable cross in the close position, triangles in the separated position. The three curves are nearly identical. Intensity-response functions for both crosses illuminated simultaneously are shown on the upper left. The circles and dashed line are for the case in which the two crosses are adjacent; the triangles and solid line are for when they are well-separated.

The response-summation and sensitivity-summation plots for each case are on the right of Fig. 15. There is no substantial difference on the response-summation plots (top) arguing for a lack of lateral interaction, but there is an apparent difference on the sensitivity-

Fig. 15. Results of a search for evidence of lateral interactions. Stimuli were four-limbed crosses (cross globicus) as shown at the top right. Each was centered on the receptive field, and they were presented either individually or together in equal-intensity experiments. One cross (shown lightly stippled) was always in the same position; the other cross could be rotated (before presentation) to either of two positions, as shown. Open circles refer to responses to the stationary cross stimulated alone; solid circles and dashed curves refer to responses when the rotated cross is in the position adjacent to the stationary cross; triangles and solid curves refer to the rotated cross separated from the stationary cross. On the lower left are the intensity-response data for the crosses presented separately; intensity-response curves for the two equal-intensity experiments are on the upper left. The response-summation plot is on the upper right (the curves for the two situations superimpose); sensitivity-summation on the lower right. Other conventions as in Figs. 7 and 8; all stimuli were 710 nm.

R55
-R+R
R_{1+2} (adrians)
60
40
20
-3.0
-2.0
-1.0
0.0
Log (I)
0.63 mm
$(R_1 + R_2)$ (adrians)
0
20
40
60
80
100
R_1 or R_2 (adrians)
$-\text{Log}\left(\frac{1}{I_1} + \frac{1}{I_2}\right)$
0.0
-1.0
-2.0
-3.0
Log (I_{1+2})

Fig. 15

summation plots (bottom). In particular, the solid curve, from the case in which the crosses are well separated, is below the dashed curve (crosses nearly touching) at low intensities. Since the effect of inhibitory interaction would be to depress this plot, the implication is that there was greater interaction when the crosses were separated than when they were tangent, and that the interaction was greater at low intensities than high.

Neither of these prospects is inviting, for one would expect interaction to increase with intensity, not decrease; moreover, one would expect interactions to be greater with smaller separations, not larger. There are two other possibilities: 1) the differences noted on the sensitivity-summation plots could be the result of a facilitation by which the adjacent crosses aided each other in overcoming the thresholds (which presumably cause the low intensity depression of the curves); or, 2) they might be due to random variance. With regard to the latter possibility, it is worth noting that the differences are at the low intensity part of the curve, which is its least reliable part; also, the entire difference can be traced to a small number of data points at the lowest intensities.

But we may not yet dismiss the possibility that the lateral influences are in fact smaller at close distances, and increase to a maximum at a slightly larger separation. This is the kind of distance function that was reported by Barlow (1967) for lateral inhibition in Limulus. We must therefore perform a similar pair of experiments in which the distances considered are somewhat larger.

A test for interactions across wider separations is shown in Fig. 16. Three 0.4 mm spots were focused on the center of this unit's receptive field. One spot (spot 1) was always used as one partner in a two spot experiment; the second area was either the one quite close to it (spot 2) or the one quite distant (spot 3). Individual intensity-response curves for each of the three spots are shown in the lower left of the figure; the spots were essentially in equisensitive positions, since the curves superimpose quite closely. The intensity-response

Fig. 16. Results of a search for evidence of lateral interactions effective over long distances. Stimuli were three spots, 0.4 mm in diameter, within the center of the receptive field. Two different two spot experiments were performed: i) spots 1 and 2 were used (open symbols, dotted curves), and ii) spots 1 and 3 were used (solid symbols, dashed curves). Peg experiments are shown by triangles; the solid curve on the intensity-response plot at the upper left fits either set of triangles. Plots arranged as in Fig. 15; other conventions as in Figs. 7 and 8; all stimuli were 710 nm.

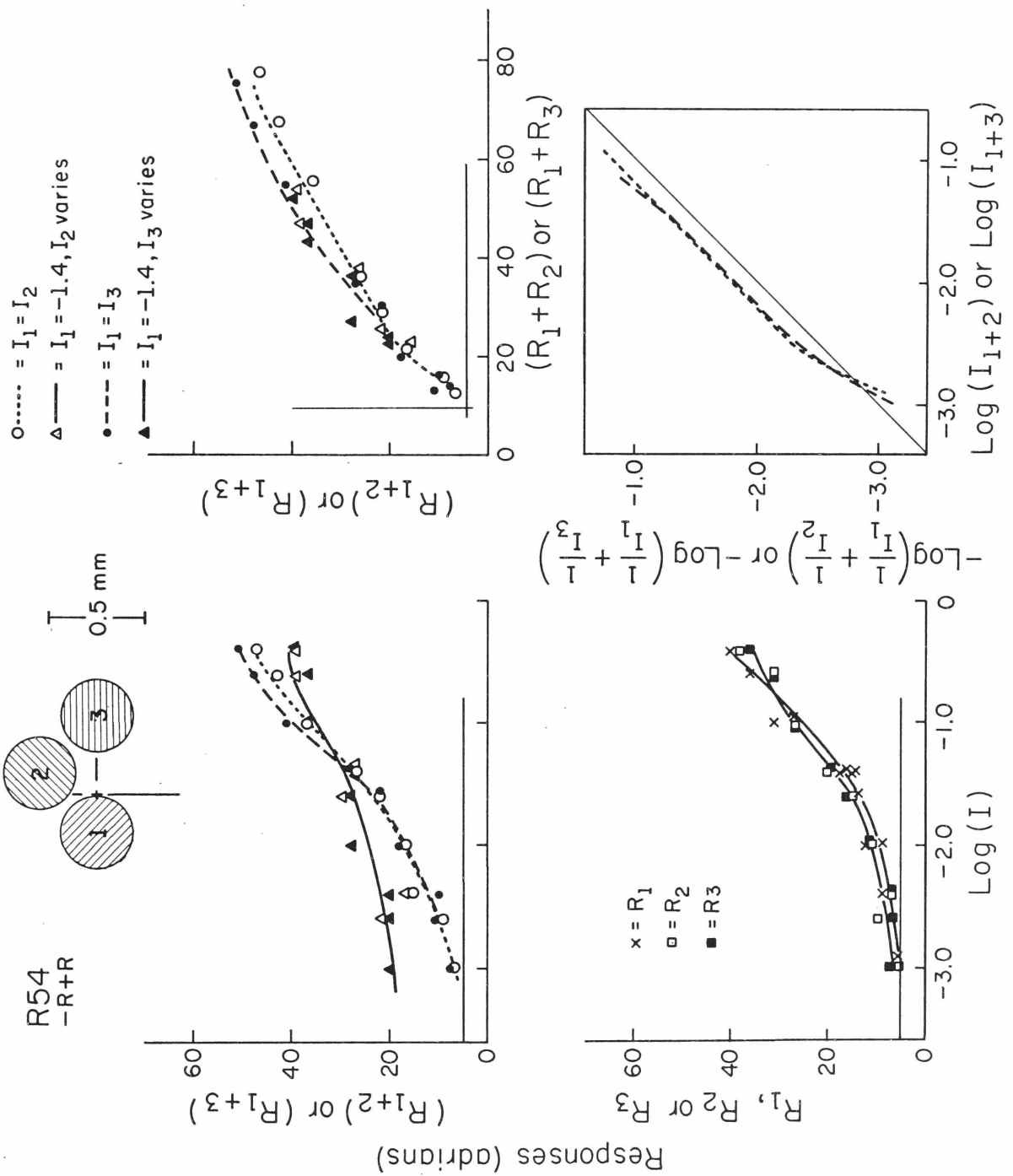
R54
−R+R
0.5 mm
1
2
3
Responses (adrians)
(R1+2) or (R1+3)
R1, R2 or R3
x = R1
□ = R2
■ = R3
Log (I)
−Log(1/I1 + 1/I2) or −Log(1/I1 + 1/I3)
Log (I1+2) or Log (I1+3)
(R1+R2) or (R1+R3)
o····· = I1 = I2
Δ—— = I1 = −1.4, I2 varies
•– – – = I1 = I3
▲—— = I1 = −1.4, I3 varies

Fig. 16

curves for the equal-intensity experiments for each pairing are shown on the upper left of Fig. 16. The equal-intensity experiment for the close spots is shown by the open circles (dotted curve), the distant pairing is shown by solid circles (dashed curve); the curves agree closely. A peg experiment was also performed for each of the pairings (triangles, solid curve); there was no difference between the outcomes of the peg experiments for the two separations.

The response-summation plots from these data are on the upper right in Fig. 16. The curves derived from the two pairings do not quite superimpose; the curve for the closer pairing falls slightly below that for the more distant pair. This is the result expected if there is more lateral interaction between the closer spots. Notice, however, that in this case the interaction is evident only at the higher firing levels; furthermore, it is not evident in the points derived from the peg experiments. Moreover, the sensitivity-summation plot, lower right, shows no substantial difference between the curves; it is probably fair to conclude that there is no evidence for lateral interaction in this unit, only variability in the data.

If there is lateral interaction its effect must be quite small, for the differences between the two response-summation curves in Figs. 15 or 16 are minor compared to the difference between any of the curves and a line of unity slope. Other cases in which different stimuli were used on the same unit show only negligible differences in the response-summation and sensitivity-summation plots for each configuration. For example, the unit whose responses are shown in Fig. 15 for two positions of the cross globicus also has its responses shown in Figs. 11 and 12 (E) for a stimulus pair consisting of a spot and surrounding annulus; the resultant curves in Figs. 11, 12, and 15 compare quite closely. As an even more extreme example, the unit shown in Figs. 11 and 12 (C) for two spots within the center of its receptive field is also shown on the right side of Fig. 14 for two areas within the surround of its receptive field. Despite this considerable difference in stimuli, the two sets of curves compare quite well.

Ricco's Law[g] Thus far I have found no consistent evidence to compel me to include lateral interactions in the model, but a final point must be considered. Easter (1968a) postulated lateral interactions to account for the differences he noted between the summation of two separate areas and the summation of contiguous areas in the experiment in which spot size was varied (see "Ganglion Cells" in the Introduction). I have already indicated an alternative explanation from Easter's own data (see "Ganglion Cells"), but I have not as yet shown that my interpretation is preferable. The data obtained from summing two areas indicate that no lateral interactions need be postulated, but can the same data be used to demonstrate Ricco's law?

The experiment shown in Fig. 17 was designed to demonstrate that the relationship of Ricco's law is consonant with the usual results of summing responses from two separate areas. Two separate pairs of stimuli were used (upper right). The first pair was a spot 0.63 mm in diameter (area 2) and an annulus of equal area whose outermost diameter was 1.0 mm (area 1); this is still just within the center of the receptive field. The second pair consisted of a spot 0.4 mm in diameter (area 4) and an annulus whose outer diameter was 0.63 mm (area 3). Intensity-response curves for areas 1 and 2 presented singly and together are shown in the upper left of Fig. 17. There was a considerable sensitivity difference between area 2 alone (spot) and area 1 alone (annulus); this difference is to be expected, since sensitivity declines with distance from the midpoint of the receptive field (see Fig. 3). The response-summation plot derived from these data is shown in the lower left of Fig. 17. The data from areas 3 and 4 are not shown, but they are very similar, and both sets of curves are typical of those found with other stimulus configurations on other units. The sensitivity-summation plot is on the lower right; the dotted curve is plotted in the usual manner, and is related to the ordinate on the left. This curve also is unexceptional, being a straight line of unity slope with an intercept of about 0.27.

Fig. 17. Demonstration that Ricco's law is consistent with the conclusions derived from the two spot experiments. Stimuli (710 nm) were two circles and two annuli, shown on the upper right; all were within the center of the receptive field. Upper left: Intensity-response curves for stimuli 1 and 2 presented separately (solid symbols and curves) and together (dashed curve, open circles); these areas are clearly not equisensitive. Lower left: Response-summation plot for stimuli 1 and 2, equal-intensity experiment. Lower right: The dashed curve, referred to the ordinate to the left, is the usual sensitivity-summation plot for stimuli 1 and 2; conventions for this curve and the graphs to the left are the same as for Figs. 7 and 8. The solid curve is derived from the same data as the dotted but is plotted according to the method of Easter (1968a), and refers to the ordinate to the right; see text for a discussion of these two forms of plotting the data. Top right: The plot shows log sensitivity vs log area for a criterion response of 30 adrians, derived from the intensity-response curves on the left. The stimuli whose areas correspond to the four values of the abscissa are indicated above each data point; the line through the points has a slope of one, demonstrating that Ricco's law holds up to the largest area used.

Responses (adrians)
R_{1+2}, R_1, R_2
• = R_2
▲ = R_1
o = R_{1+2}
Log I
R_{1+2}
$(R_1 + R_2)$ (adrians)
R54
-R+R
Log sensitivity
Log area
1+2
3+4
2
4
0.5mm
1
2
3
4
$-\text{Log}\left(\frac{1}{I_1}+\frac{1}{I_2}\right)$
----- = $-\text{Log}\left(\frac{1}{I_1}+\frac{1}{I_2}\right)$
—— = Log (I_2)
Log (I_{1+2})
Log (I_2)

Fig. 17

Now let us examine how these same data would appear if they were plotted exactly as described by Easter (1968a). We would assume that areas 1 and 2 were still within the plateau of equal sensitivity, and make use only of the intensity-response curves for area 2 alone and for both together. (When changing spot sizes, the intensity-response curve for the additional contiguous annulus would not be measured separately.) When the data are analyzed in this fashion, the sensitivity-summation curve is the solid curve whose ordinate is indicated on the right. This ordinate is displaced by 0.3 log units (see "Method of Sensitivity-Summation"), so that the implications of the line drawn as the diagonal are the same as for the usual sensitivity-summation plot; that this curve lies along the diagonal implies complete linear summation and linear receptors. Easter, having shown the receptors were non-linear by summing well-separated spots, would have assumed lateral interactions account for the location of the curve. But both the solid curve and the dotted curve were derived from the same set of data, so the difference between them must be due to the method of analysis, and not the presence or absence of interactions.

Finally, it is necessary to show that Ricco's law was demonstrated by these data. The sensitivity of each of the four areas used was derived from the individual intensity-response curves using a criterion response of 30 adrians. The log sensitivity of each area was plotted as a function of the log of the area (Fig. 17, upper right) and a line of unity slope drawn through the points. If S is sensitivity, A is area, I is intensity, and k and k' are constants:

$$\log S = \log A + k$$

or

$$-\log I = \log A + k$$

$$A \cdot I = k'$$

which is Ricco's law. The same data that gave unexceptional response-summation and sensitivity-summation plots were used to demonstrate Ricco's law. There is no contradiction requiring the postulation of

lateral interactions for its resolution; as discussed earlier ("Ganglion Cells" in the Introduction), it is sufficient to have a square root function at the receptors and to have the sensitivity across the center of the receptive field fall off as the inverse square of the distance from the midpoint.

Response Segments

All of the analyses up to this point have been based on responses averaged over a full second, but the firing of a ganglion cell is far from uniform during this time span. If we consider only "on" responses, we observe that the response is quite vigorous at the onset of the light, but decays to nearly spontaneous rate as the light remains on (see Fig. 5). This is shown in Fig. 18, in which the firing rate of the ganglion cell is plotted as a function of time. The stimuli were two spots within the center of the receptive field which were simultaneously illuminated at each of three intensities (equal-intensity experiment). The responses of this unit may be analysed in the usual manner, using the average firing rate over the full second the stimulus was on; such an analysis of this unit is shown in the response-summation and sensitivity-summation plots on the left of Fig. 19. (The records of Fig. 18 contributed the points at R_{1+2} = 42, 49, and 59 adrians on the response-summation plot at the top left.) But the response may also be taken as the rate during any portion of the stimulus time; three such segments are indicated by arrows on the representative records of Fig. 18. The three segments chosen were 100 to 300 msec from onset, 300 to 500 msec, and 500 to 1000 msec. The first of these segments encompassed the peak of the firing rate, the second segment was during the initial part of the decay in firing rate, and the final segment was when the rate had settled to a lower value. A full set of intensity-response curves was found using each of the response segments indicated above; from each of these, a response-summation plot and a sensitivity-summation plot was derived. The three resultant response-summation plots are shown on the same axes in the upper right of Fig. 19; the superimposed sensitivity-summation plots are on the lower right.

Fig. 18. Responses of a unit to simultaneous stimulation by two spots of light (0.4 mm diameter, 710 nm) in the center of its receptive field. Responses to three different intensities are shown (equal-intensity experiments): 0.0, -0.6, and -1.0 log units. Stimuli were 1 second flashes, indicated by a thickening of the abscissa and vertical lines at onset and offset. Frequency of firing was measured in adrians for 50 msec time periods (details of computation given by Schoenfeld, 1964).

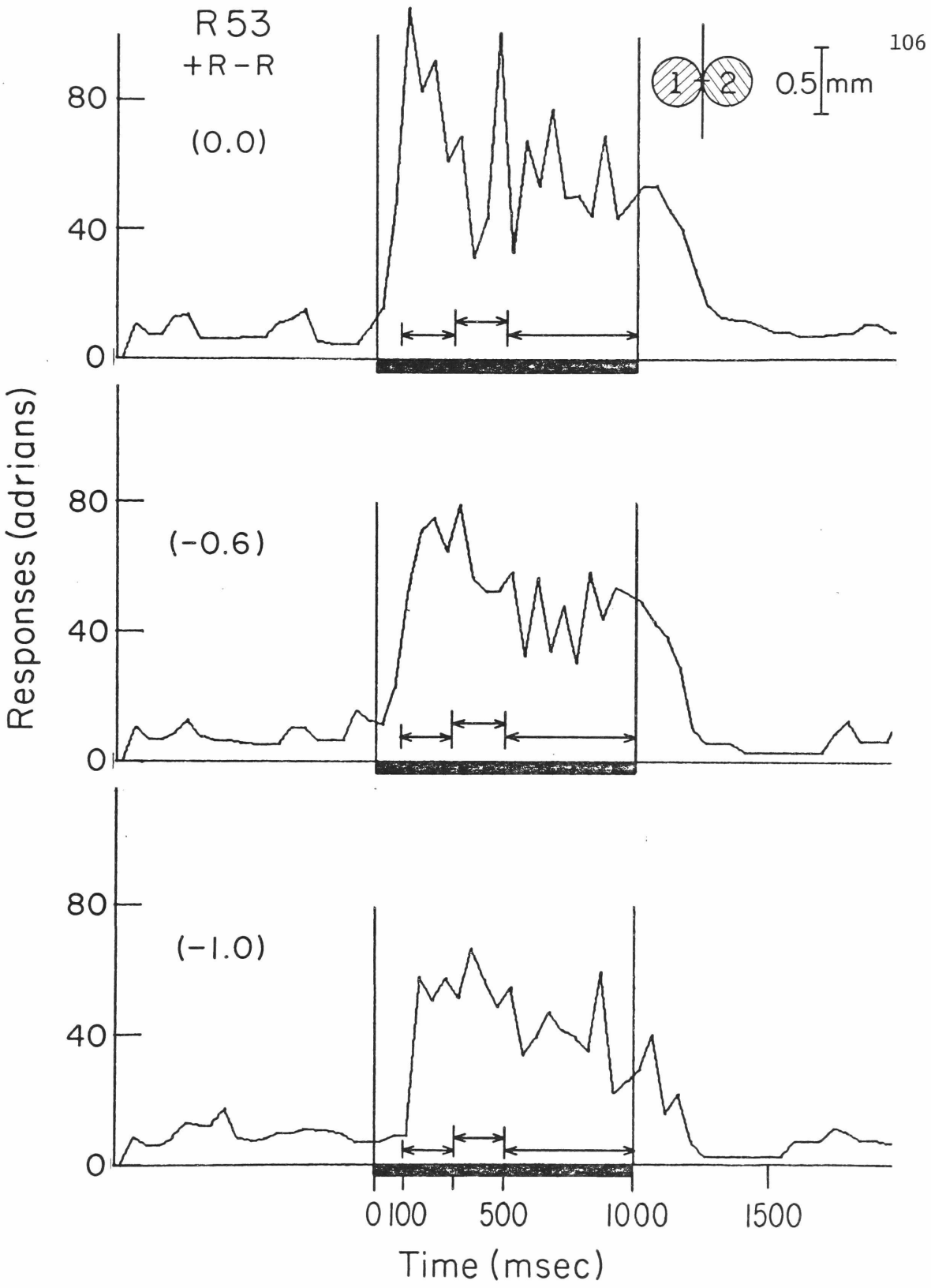
R 53
+R-R
(0.0)
(-0.6)
(-1.0)
80
40
0
1
2
0.5 mm
Responses (adrians)
0 100 500 1000 1500
Time (msec)

Fig. 18

Fig. 19. Response segments analysis of the unit of Fig. 18. Left: Response-summation plot (top) and sensitivity-summation plot (bottom) for responses averaged over the full second of stimulation. Conventions as in Figs. 7 and 8. Right: Response-summation plots (top) and sensitivity-summation plots (bottom) for three different periods within the second of stimulation; 100 to 300 msec from onset of the stimulus, 300 to 500 msec from onset, and 500 to 1000 msec from onset. These periods are indicated on the graphs in Fig. 18.

Total responses
R53
+R-R
0.5 mm
1+2
R_{1+2} (adrians)
Response segments
msec
○ 100-300
△ 300-500
• 500-1000
$(R_1 + R_2)$ (adrians)
$-\log\left(\frac{1}{I_1}+\frac{1}{I_2}\right)$
Log (I_{1+2})
Log (I_{1+2})

Fig. 19

The three response-summation plots coincide closely at low levels, where each is a straight line of unity slope, but at high levels they are clearly separate curves. The differences among the three may be approximately described as one of scale; any of the three can be superimposed on another by the appropriate multiplication of both ordinate and abscissa by the same factor. This same observation was made on two other units from which "on" responses were obtained from two areas in the surround of the receptive field.

At some stage in the retinal processing the signal is dependent on the history of the illumination of the retina. One might describe the observed time course of the response of the ganglion cell by postulating a stage of differentiation, or high-pass filtering, so that signals which are changing in time are amplified, while steady signals are attenuated. For the purposes of the model shown in Fig. 13, we can state simply that there is a time dependent factor*, F(t), that multiplies the signal. Let us consider where this factor might appear in the model so that it will result in the contraction of the response-summation plots noted above.

If F(t) were a property of the receptors, the multiplier could be considered part of the receptor function, and would therefore not affect the response-summation plot at all (see discussion of v in Appendix II); similarly, a multiplication after summation but before the final compressive function would be mathematically equivalent to multiplication before summation, and undetected. With either of these multipliers, the effect would be not to change the curve, but to trace different segments of the same function by changing the range of the argument.

---

*Since all the stimuli used in these experiments were steps of light, I shall use F(t) to mean the time dependent component of the response to a step; that is, the convolution of the impulse response, $F_\delta(t)$ with a unit step change in input. The most general way to characterize the time function would be the transform of the impulse response into the frequency domain: f(s), as discussed in Appendix IV.

The only place such a multiplication could take place so that the response-summation plots for the different segments would fall on different curves would be after the final compressive function (see discussion of changing g, Appendix II). The expected result of changing a multiplicative factor at the output would be a uniform contraction or dilation of the response-summation plot, which is exactly what was observed.

No matter where in the model we place a multiplicative F(t), it should make no difference to the sensitivity-summation plots (see Appendix III), but the sensitivity-summation plots on the lower right of Fig. 19 do appear to show differences. In particular, the plot derived from the segment covering the peak firing rates indicates a higher exponent for the receptor function than is indicated by plots from later segments. It is possible that the receptors are in fact somewhat more linear in the initial portion of their responses, but these data are not sufficiently compelling to warrant such a conclusion. The differences in the sensitivity-summation plots for different time segments were not corroborated by the analyses of other units.

Relative Timing

I have concluded that the time dependent multiplier is late in the system, which is consistent with the findings discussed in the Introduction ("Inner Nuclear Layer") that units distal[g] to the amacrine cells respond tonically (e.g., Kaneko, 1970). But other experiments raise some questions about this interpretation; also, the rather broad conclusions I have presented are confined to only the center of the receptive field (or only the surround). The following experiments, referred to as "relative timing," suggest how these findings might be extended to include interactions of center and surround. The reader is referred back to the discussion of relative timing under "Experimental Paradigms" for a description of the experiment.

Fig. 20 shows the results of a relative timing experiment in which the test flash (200 msec) was in the center of the receptive field of the ganglion cell, and the conditioning flash (1 second) was in the surround. Responses to test flashes in the absence of any conditioning flash are shown on the graph marked "T" on the left; the stimulus lasted for the period from 0 to 200 msec (solid vertical lines), and the firing rates shown are the average for each 50 msec interval. Three different stimulus intensities were used, but no substantial differences in the time course of the responses are apparent. For purposes of analysis, the response to the test flash was therefore taken as the average firing in the period from 350 to 500 msec from the onset of the test flash; this period is indicated by the vertical dashed lines.

When both test and conditioning lights were presented, the response is plotted on a time scale for which the onset of the conditioning stimulus is defined as 0 (upper right of Fig. 20). Response was measured in the period from 350 to 500 msec after the test flash onset, and plotted according to when it occurred relative to the conditioning stimulus. For example, a symbol plotted just to the right of the vertical line (at 0.025 sec) in the upper right graph represents the responses measured in a period from 50 msec before the conditioning stimulus to 100 msec after the onset of conditioning stimulus; the test flash in the record from which this point was derived covered the time from 400 msec to 200 msec before the conditioning stimulus. Each point on this graph is derived from a single record, and represents the total response to the test flash in the presence of the conditioning stimulus. The relative timing of test and conditioning were varied, and three different intensities of test stimulus were used.

The graph in the center of the right column of Fig. 20 shows the response to the conditioning stimulus in the absence of a test flash. Each point represents the average rate during the 150 msec period of which it is the center; this is the contribution of the conditioning stimulus to total response at that moment in time.

Fig. 20. Relative timing experiment; stimuli (shown at upper left) were a 1 second conditioning annulus (C) whose intensity was -1.6 log units, and a 200 msec test spot (T) that could have one of three intensities (-0.6, -1.0, and -1.6 log units). All stimuli were 710 nm. Left: Responses to the test flash, measured as number of spikes in successive 50 msec periods (post stimulus time histogram) and converted to adrians. Stimulus onset (time = 0) and offset (time = 200 msec) indicated by solid vertical lines. The upper graph shows responses to the test flash at each of three different intensities presented in absence of the conditioning stimulus. Lower graph shows responses to the test flash at -1.0 log units, presented at different times (α, β, γ,δ) relative to the conditioning stimulus; response is taken as number of spikes in each 50 msec period less the number of spikes in the corresponding period for the response to the conditioning stimulus in the absence of a test flash. For analysis, responses in the period (indicated by dashed vertical lines) from 350 to 500 msec after onset of the test flash were taken as the response to the test flash. Right: Responses averaged over 150 msec periods and converted to adrians. Time along the abscissa measured from onset of conditioning stimulus (vertical line at time = 0). Top graph shows responses to the test flash in the presence of the conditioning stimulus; responses were counted for the period from 350 to 500 msec from onset of a given test flash. Middle graph gives response (150 msec periods) to conditioning stimulus in absence of any test flash. Bottom graph shows the contribution of the test flash to total response, found by subtracting the middle graph from each of the curves in the top graph. The points marked α, β, γ, and δ are the times for which the responses are shown on the graph at the lower left.

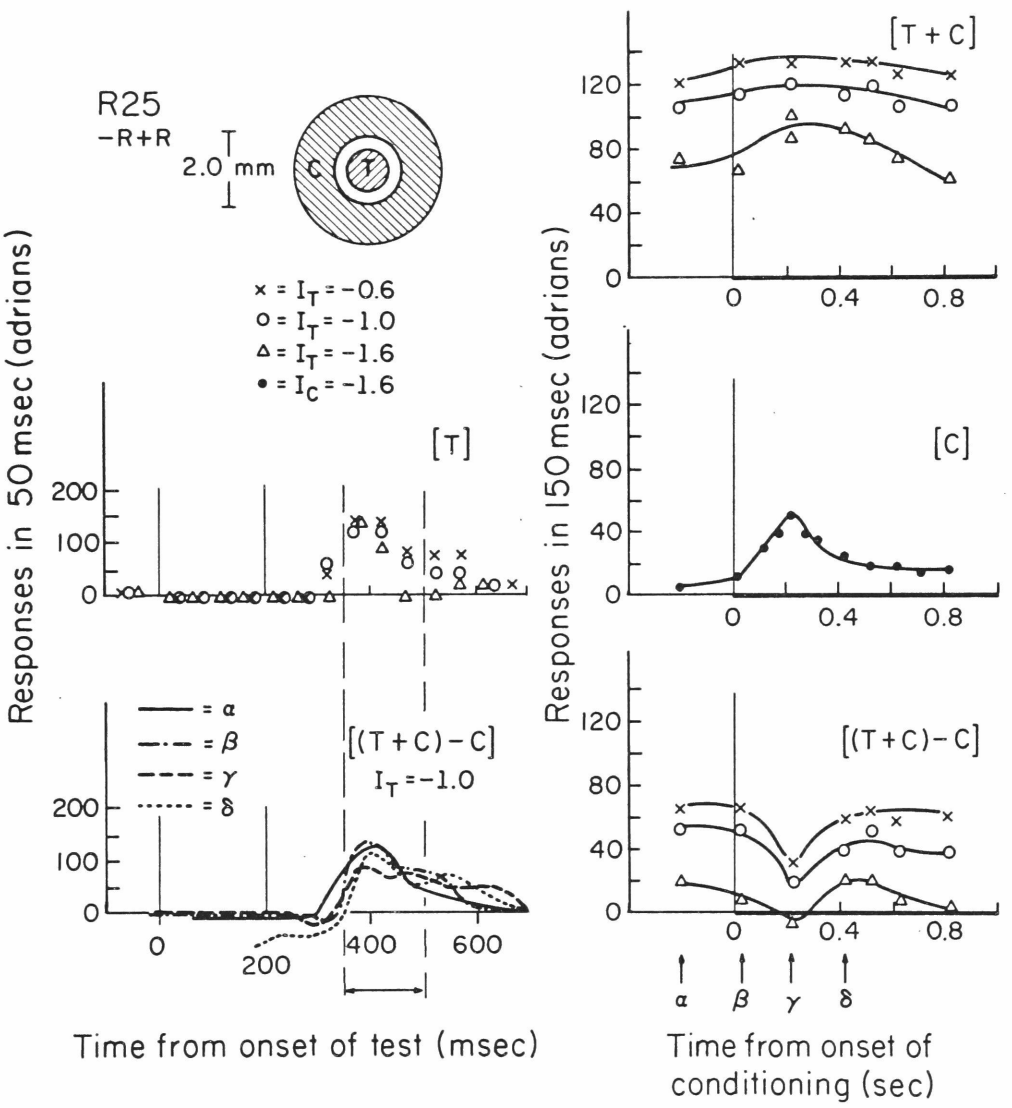
R25
-R+R
2.0 mm
C
T
× = I_T = -0.6
○ = I_T = -1.0
△ = I_T = -1.6
● = I_C = -1.6
[T]
[C]
[T + C]
[(T + C) - C]
I_T = -1.0
α = α
β = β
γ = γ
δ = δ
Responses in 50 msec (adrians)
Responses in 150 msec (adrians)
Time from onset of test (msec)
Time from onset of conditioning (sec)

Fig. 20

Since the top graph is the response to the test flash plus the conditioning stimulus, and the middle graph is the response to the conditioning stimulus alone, a point by point subtraction of these two curves should give that portion of the response due to the test flash. If the response to the test flash were independent of the conditioning stimulus, a straight, horizontal line would result. The actual result of this subtraction is shown on the lower right of Fig. 20; there is a distinct dip in the curve, corresponding in time to the peak of the response to the conditioning stimulus. This finding is consonant either with lateral interactions between the areas stimulated by the test and the conditioning stimuli, or a non-linear (compressive) function after the summation of the two areas.

But it is also possible that the dip is an artifact of the time period chosen for counting responses; this might occur if there were a change in latency of the response to the test flash as a consequence of the presence of the conditioning stimulus. To test this hypothesis, I have found the responses to a -1.0 log intensity test flash presented at various times (α, β, γ, δ) relative to the conditioning flash: response to the test flash is found by subtracting the responses to the conditioning stimulus alone from the responses to test plus conditioning stimuli. The results are shown in the lower left of Fig. 20; α, β, γ, and δ are indicated on the lower right abscissa. There is no consistent change in the shape of the response to the test flash, only a change in amplitude.

The relative timing experiment was performed on three other cells, as shown in Fig. 21. The unit shown on the left, like the unit in Fig. 20, had the conditioning stimulus in the surround of its receptive field and the test flash in the center. The unit in the center column of Fig. 21 had both the test and conditioning lights in the center of its receptive field, while the unit shown on the right had both stimuli in the surround of its receptive field. The results are essentially the same in all cases; the contribution of the test flash found by subtracting the response to conditioning alone from response to test plus conditioning

Fig. 21. Relative timing experiments on three units; stimuli sketched at the top for each. Conventions are the same as for the right side of Fig. 20. Left: Test stimulus (-2.0 log units) in the center of the receptive field, conditioning stimulus (-2.6 log units) in the surround. Center: Both test flash (-1.4 log units) and conditioning stimulus (-1.4 log units) were in the center of the receptive field. Right: Both the test flash (-1.6 log units) and the conditioning stimulus (-1.6 log units) were within the surround of the receptive field. All stimuli were 710 nm.

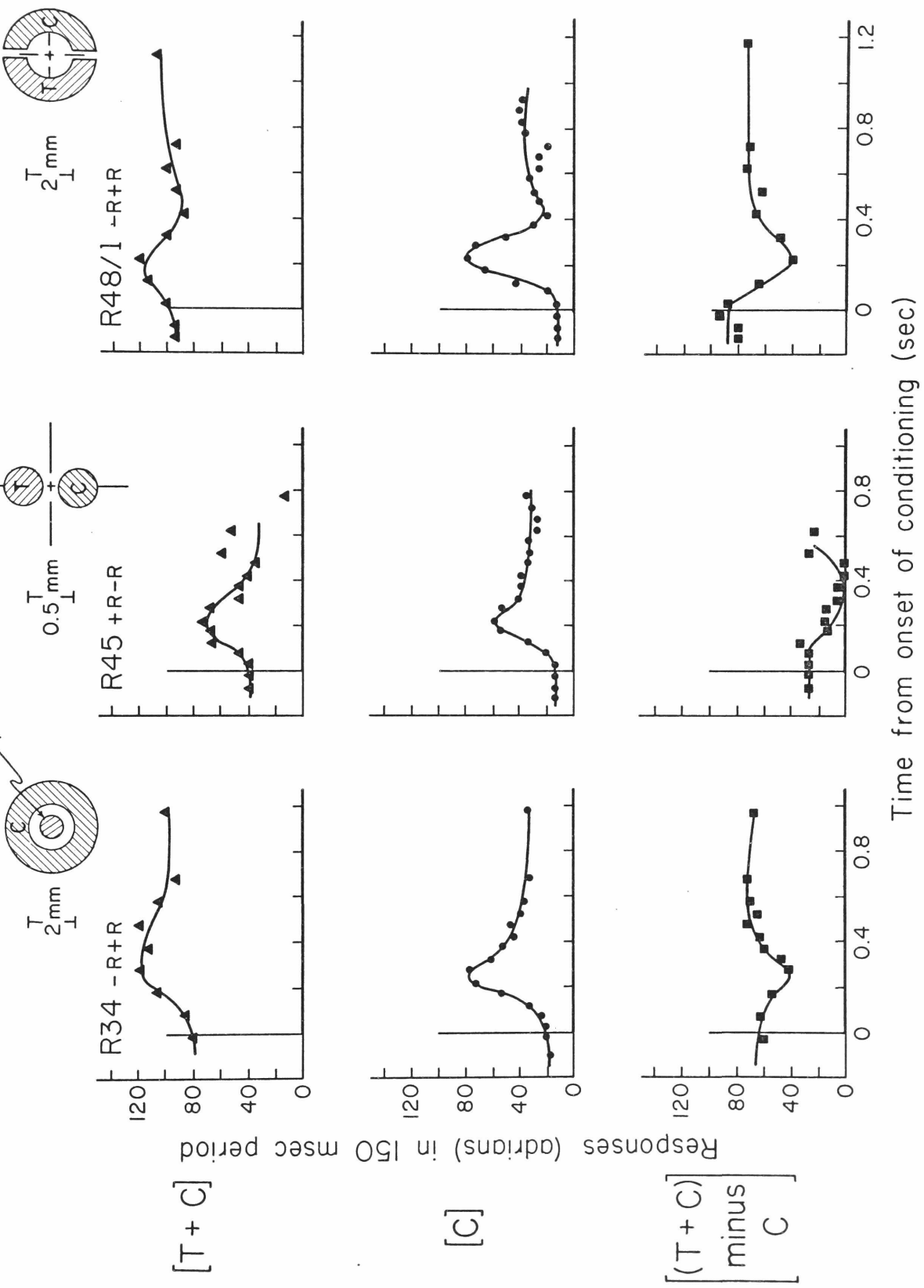
R34 -R+R
R45 +R-R
R48/1 -R+R
T 2mm
T 0.5mm
T 2mm
C
T
[T + C]
[C]
[(T + C) minus C]
Responses (adrians) in 150 msec period
Time from onset of conditioning (sec)
0
40
80
120
0.4
0.8
1.2

Fig. 21

shows a dip roughly coincident with the peak firing of the response to the conditioning stimulus alone.

At first, this result may seem to contradict some of the conclusions from the "response segments" analysis (see "Response Segments"), from which I concluded that the change in response rate with time is due to a time dependent multiplier, F(t), after the compressive function after summation. Here, it seems that the change in response to the test flash depended on the time course of the response to the conditioning stimulus, so its temporal properties must have been expressed before the compression. This paradox may be resolved by a more careful consideration of how F(t) is determined.

Let us assume that the compressive function after summation is the type equivalent to an inclusive forward shunting inhibition (see "Interaction vs Compression After Summation"), equation (2):

$$R = \frac{(x_1 + x_2)}{k + (x_1 + x_2)} = \frac{s}{k + s}$$

Assume further that there is an F(t) on each input to the summation point, y_i (see Fig. 13), and that each F(t) depends only on the time t_i since the onset of that input. If we consider only responses to a step of light, then:

$$F_i(t) = F(t_i) = f \qquad (3)$$

Let us assume that the instantaneous values of this function (f) multiply only the numerators in equation (2); in other words, that the time dependence is only in the straight-through pathways of the model shown in Fig. 13, not the divisors:

$$R = \frac{g[F(t_1) \cdot y_1 + F(t_2) \cdot y_2]}{k + (y_1 + y_2)} \qquad (4)$$

(Subscripts refer to separate stimulus areas; other symbols used here are the same as in Appendix II.)

In the two spot experiments, $t_1 = t_2$, because the two stimuli are simultaneous; therefore equation (4) becomes:

$$R = \frac{g \cdot F(t) \cdot (y_1 + y_2)}{k + (y_1 + y_2)} = \frac{g \cdot F(t) \cdot s}{k + s} \qquad (5)$$

which is the result expected from the response segments analysis.

In the relative timing experiment, we are always measuring response rates at a fixed time after the test flash, so $F(t_2)$ always has the same value: call it f_2. (The subscript 1 now refers specifically to the conditioning stimulus, 2 to the test.) The value of $F(t_1)$, however, depends on which time period after the onset of the conditioning stimulus is being considered; assume that at the peak of the response to the conditioning stimulus it has some value f_1, while much later in the response its value is negligible. The response to the test flash in the absence of a conditioning stimulus is:

$$R_T = \frac{g \cdot f_2 \cdot y_2}{k + y_2} \qquad (6)$$

The response to the test in the presence of a conditioning stimulus is given by:

$$R_{T+C} = \frac{g[f_2 y_2 + F(t_1) \cdot y_1]}{k + y_1 + y_2} \qquad (7)$$

The response to conditioning stimulus alone is:

$$R_C = \frac{g \cdot F(t_1) \cdot y_1}{k + y_1} \qquad (8)$$

The contribution of the test flash to the total response is obtained by subtracting R_C from R_{T+C}; this result may be compared to R_T. If we first consider a test flash that occurs well after the onset of the conditioning stimulus, we can predict the results of the subtraction with the substitution $F(t_1) = 0$ in equations (7) and (8):

$$R_{T+C} - R_C = \frac{g \cdot f_2 \cdot y_2}{k + y_1 + y_2} \qquad (9)$$

The contribution of the test flash found by equation (9) is thus slightly smaller than the response to the test flash alone given by equation (6). Now let us consider the case in which the response to the test flash falls on the peak of the response to the conditioning stimulus. Subtracting equation (8) from (7), and substituting $F(t_1) = f_1$:

$$R_{T+C} - R_C = \frac{g \cdot f_1 \cdot y_1}{k+y_1+y_2} - \frac{g \cdot f_1 \cdot y_1}{k + y_1} + \frac{g \cdot f_2 \cdot y_2}{k+y_1+y_2} \qquad (10)$$

The third term in equation (10) is the same as the right hand side of equation (9); since the second term in (10) is larger than the first, the value of the response to the test flash found in equation (10) for a response coinciding with the peak of the response to the conditioning stimulus will be less than the value of the contribution of the test flash found from equation (9) for a response well after the peak.

One further comment may be made with respect to the results of the relative timing experiment. Fig. 21 shows experiments in which both test and conditioning were in either the center or the surround of the receptive field, and an experiment in which the test flash was in the center and the conditioning stimulus in the surround; in all cases, only excitatory responses were considered. There are no substantial differences in the results for these three cases; there is therefore some indication that the way in which excitatory influences interact are the same regardless of where in the receptive field they arise. The implication is that the excitatory part of the response from the periphery sums with the excitatory part of the response from the center at a point mathematically equivalent to the point at which sub-areas of the center are summed. Notice that an excitatory influence from the surround and an excitatory influence from the center would be of opposite type; I am referring to the summation of an "on" response and an "off" response. This makes no statement about the manner in which the inhibitory influences are summed or interact with the excitatory influences.

I have not attempted to state the configurations of the center and surround of the receptive fields. It is quite possible that, as is the case in the cat (Rodieck and Stone, 1965), the center and surround each cover the entire receptive field. In that case, all the stimuli I have used would have stimulated both the center mechanism and the surround mechanism, the type of response being determined by which was the more sensitive in that location. Every response would thus be the result of the interaction of an excitatory and an inhibitory influence from the identical area; how these influences are combined might determine whether or not my results would have been affected by this situation.

If the two influences combine linearly (subtract), this would have no effect at all: if the excitation and inhibition sum at the same point as the sub-areas of the center, the effect would be the same as a change of scale of the receptor function (v); if they sum after the compressive function (F), the effect would be the same as a change of scale of the compressive function (g) (see Fig. 13). On the other hand, a division of center by surround (as suggested by Sperling, 1970) would result in a compressive function. If the division occurred before the summation of sub-areas, it would appear as an effect on the receptor function; the input to the summation would be (see Fig. 13):

$$y_i = \frac{x_i}{1.0 + k \cdot x_i} \qquad (11)$$

where k is a constant, and x_i and y_i are defined in the usual way. If the function at the receptors is a power relationship:

$$e(I) = I^n = x_i$$

then equation (11) is of the form suggested by Boynton and Whitten (1970), and the effect of the division would be to cause the receptor function to tend toward a ceiling at high intensities; this might account for the upturn occasionally observed on the sensitivity-summation plots (see "Sensitivity-Summation Plots" in Experimental Results and Conclusions). If the division occurred after summation, it would have the effect of

inserting an additional serial compressive function, which is the same as changing the constants in the postulated compressive function (see "Compressive Function After Summation" in Appendix II). It thus appears that an overlapping of center and surround, of some plausible form, would not have had a qualitative effect on the results I have reported.

SUMMARY

The principle purpose of this thesis was to develop a minimal model for the processing underlying the receptive field of a ganglion cell. In order to achieve this, it was necessary to develop new analytical tools; it is hoped that these analytical methods will prove as valuable in future work as they have in this study.

Methods of Analyzing Two Spot Experiments

Two coordinate systems were developed for the presentation of the data from experiments in which two separate areas within the center of the receptive field were stimulated. The response-summation plot (physiological sum vs algebraic sum) is affected by non-linearities after the first interaction of the two areas; the sensitivity-summation plot (sum of sensitivities of each spot alone vs the sensitivity of both stimulated simultaneously) is affected by non-linearities before the final interaction of the two areas. By comparing the two plots, effects can be localized to one of three regions: non-linearities before any interaction affect only the sensitivity-summation plot, non-linearities after the final interaction affect only the response-summation plot, while non-linearities between the first and the final interaction affect both plots.

This form of analysis can be extended to any situation in which two separable inputs each contribute to a single output. The two stimuli need not be two separate areas in the receptive field; they might be cones containing two different pigments, but because both types of cone may be found in the same area of retina there is the difficulty of stimulating cones containing a short-wavelength sensitive pigment to the exclusion of cones containing the long-wavelength pigment. Nor is this analysis limited to situations in which two separate spots of light are the stimuli; it might be applied to any sense modality, or to a situation in which two neurons converge ultimately onto a third and all three are recorded.

I do not mean to imply that the method is without limitations. The sensitivity-summation plot cannot be generated directly unless the influences from the two stimuli are either both excitatory or both inhibitory; this is because a given criterion response would never be achieved by each stimulus separately. It is possible, however, that this difficulty may be circumvented by embellishments of the two spot experiments making use of "on" and "off" responses from the two stimuli.

Model of the Center of the Receptive Field

The application of these analysis methods has led to a minimal model of the interactions between groups of cones containing the long-wavelength sensitive pigment located entirely within the center (or entirely within the surround) of the receptive field of ganglion cells in goldfish.' This model, shown in Fig. 13, is both self-consistent and compatible with data from other workers (see "Ganglion Cells" in the Introduction).

The model depicts the signals from the two separate areas being summed linearly; in addition, there are non-linear transformations of each signal before the two are summed, and a non-linear transformation of the combined signals. The non-linearity before summation is probably a property of the receptors themselves, and may best be described as a square root relationship between light intensity and excitation. The square root relationship fails at extremes of illumination: at very low intensities, the receptors act linearly; at high intensities, receptors reach a limit of maximum possible excitation. Excitations must exceed some threshold level to be included in the sum.

The non-linearity after summation may best be described as a compressive function: at low levels, the firing rate is linear with the sum of the excitations; at high levels, the firing asymptotically approaches a maximum. It is at this late level of processing that the changes in nervous activity during a maintained stimulus occur. The two spot experiments indicate that these temporal changes could occur after the compression, but the relative timing experiments show that both the

compression and temporal functions probably occur at the same level. One plausible scheme would be to have a direct pathway which includes a temporally dependent multiplier, and an indirect pathway which sums all the excitations but does not have the temporal multiplier. The output of the system would be the result of dividing the direct pathway by a constant plus the indirect pathway (see "Relative Timing" in Experimental Results and Conclusions).

This model does not include any lateral interactions, as distinguished from the summation of the sub-areas of the receptive field. This does not necessarily mean that there is only one level of interactions between the sub-areas; I have suggested that the compressive function after summation might well have been generated by two independent summations, and we have seen that a completely linear interaction would be undetectable by the methods used here. The point is that this model, without interactions, is a sufficient description of the summation of the two areas.

The classic example of lateral interaction is the lateral inhibition described for the horseshoe crab, Limulus by Hartline (1949; see also Ratliff, 1965; Hartline, 1969). In that case, the ommatidium being recorded is excitatory, but is surrounded by a ring of inhibition via lateral pathways from all other illuminated ommatidia. Lateral inhibition in Limulus might better be compared to the inhibitory surround of the vertebrate ganglion cell's receptive field. The psychophysical phenomena such as accentuation of edges and Mach bands which are attributed to lateral inhibition (Ratliff, 1965) have been attributed to the inhibitory surround in the receptive fields of ganglion cells of the cat (Enroth-Cugell and Robson, 1966; Rodieck and Stone, 1965; Baumgartner, 1961). Similarly, in frog, Gordon (1972) has found that there is no edge accentuation of stimuli confined entirely within the center of the receptive field of the ganglion cell from which he recorded. In recordings from lateral geniculate cells in the macaque monkey, De Valois and Pease (1971) found evidence for accentuation of luminance boundaries,

but no accentuation of boundaries of different colors equated for luminance; simultaneous color contrast has been interpreted in terms of lateral interactions at boundaries, but in primates this does not seem to take place in the retina.

The model I have presented is a sufficient description for the simplified case in which spatial antagonisms have been eliminated by stimulating entirely within the center or entirely within the surround of the receptive field, and in which chromatic interactions have been eliminated by stimulating with a near infrared stimulus effective only for the cones containing the long-wavelength pigment. Clearly, the observations must be extended to include these antagonisms, so that all the mechanisms contributing to the response of a ganglion cell may be included in the model.

APPENDIX I: GLOSSARY AND SYMBOL TABLE

Glossary

Adrian Unit of nerve firing rate: 1 adrian = 1 action potential per second.

Bloody Hell! See "Shit."

Cell The term cell is used to denote any of the neuronal cells in the retina; it is most often used to refer to the particular ganglion cell from which spikes are being recorded (see "unit"). There are a variety of different cell types in the retina; see Fig. 22 for a schematic diagram showing the various cells found in the goldfish retina.

Compressive Functions A compressive function in general is a function whose output does not increase proportionately with its input, but is negatively accelerated. This broad definition encompasses such functions as logarithms and power functions with exponents less than 1; I use the term more specifically to mean a function that is approximately linear for small values of its argument but departs from linearity at high values, approaching an asymptote which it can never exceed. The particular algebraic form that I use as an exemplar is

$$F(\zeta) = \frac{g \cdot \zeta}{k + \zeta}$$

where g and k are constants.

Difference Spectrum If the absorption spectrum of a photopigment is measured, the optical density at each wavelength is determined by the sum of the optical densities of the photopigment and all the non-photolabile material present, such as parts of the receptor and the inactive portions of the pigment molecules. (Density is a logarithmic measure, hence there is summation, not multiplication.) To correct for this, one takes advantage of the fact that the photopigments are photolabile, and are bleached away by light. The technique is to measure the absorption spectrum of the fresh pigment, then illuminate it with a light intense enough to bleach all the photopigment, and finally remeasure the absorption

Fig. 22. Schematic of the anatomy of the goldfish retina, showing layers approximately to scale. Pigment epithelium not shown. Receptors (rods and cones) synapse with bipolar cells and two types of horizontal cells at the outer plexiform layer. Bipolar cells synapse primarily with amacrine cells, which synapse extensively with each other before synapsing with ganglion cells; this processing occurs in the inner plexiform layer. Ganglion cell axons course along the inner surface of the retina toward the optic disc, at which point they pass out of the eyeball as a large bundle (optic nerve) which goes to the optic tecta. This drawing is a composite from a number of workers. General features are based on Dowling (1970) and De Test[a] (1966); scaling of the layers is from histology by Thorpe (1971). Features of the receptors are primarily from Walls (1942); the outer plexiform layer and horizontal cells are based on the work of Stell (1967). The inner plexiform layer is based on findings in other lower vertebrates by Dubin (1970).

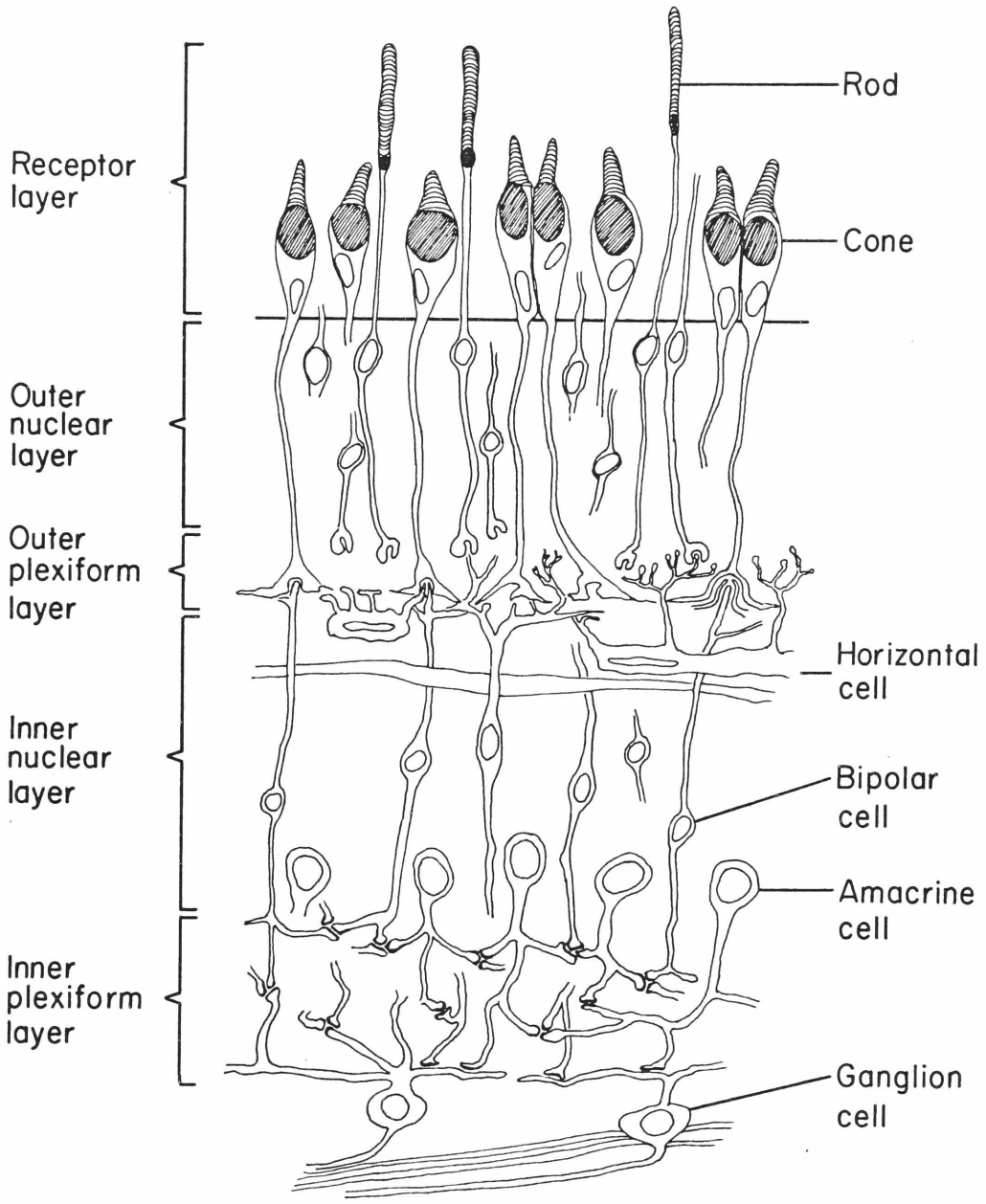
Receptor layer
Outer nuclear layer
Outer plexiform layer
Inner nuclear layer
Inner plexiform layer
Rod
Cone
Horizontal cell
Bipolar cell
Amacrine cell
Ganglion cell

Fig. 22

spectrum. The ratio of the two measurements (density difference) must be due only to the properties of the photopigment, since all other photostable absorptions appear in both terms and cancel (see Dartnall, 1962).

Distal The term distal refers to the direction in the retina away from the brain (in terms of which units serve as inputs to other units, not physical location, which is the reverse). Thus, the most distal units of the retina are the receptors, and the most proximal units are the ganglion cells.

Equal-Intensity Experiment A two spot experiment in which the intensities of the stimuli in each area are equal. See "Experimental Paradigms" in Materials and Methods.

Intensity Intensity is used in its most colloquial sense, the light flux in a stimulus. Since all the experiments reported in this thesis were done at a single wavelength (710 nm), there is no need to specify a spectrum, and the intensity can be understood to mean either the number of quanta incident on the stimulated part of the retina per unit area per unit time, or the power per unit area.

Intensity-Response Plot A plot of the responses of a unit (along the ordinate) as a function of the logarithm of the intensity of the stimulus (abscissa). The stimulus is always of the same shape, size, and wavelength, and in the same position; only its intensity is varied.

Lateral Interactions These interactions refer to an influence by which the action of light in one part of the retina can affect the responses from another area. I do not mean to include the summation of responses from different areas of the receptive field, but intend for interactions to refer to a separate effect distal to the summation of areas.

Although I do not specify, I generally mean lateral interactions to be inhibitory; the influence of one illuminated area on another should be in the direction of decreasing the response to stimulation of the other area. There is also the possibility of lateral facilitation,

in which one area might enhance the responses of its neighbor. In the text, I consider various forms of interaction in order to decide whether or not there are interactions present; in general, the interactions I consider are inhibitory. However, if the interactions were in fact facilitatory, the arguments would not be invalid; the predicted effects would merely be in the opposite direction. Thus, if the postulated inhibition would be expected to depress a particular curve, and the data show no such depression, facilitation may also be ruled out if the curve shows no elevation.

Nomogram Specifically, I am referring to a photopigment nomogram. A nomogram is a convenient graphical method for finding the spectral absorption of a photopigment given only the wavelength of maximum absorption. If the spectral absorption (on a quantum basis) of a visual pigment is plotted as a function of wave number (inverse of wavelength), a curve with a characteristic shape is obtained. Most visual photopigments whose spectral characteristics have been studied have absorption spectra of the identical shape; the curve is merely translated along the wave number axis to account for different peak absorptions. Nomograms have been constructed which allow a simple graphical solution for a pigment's absorption at all wavelengths once the wavelength of maximal absorption is known. The best known of these is the Dartnall (1953, 1962) nomogram, which is a close fit to all known visual pigments which are based on retinene$_1$ (e.g. rhodopsin). However, in many fish the pigments are based on retinene$_2$ (e.g. porphyropsin), and a slightly different nomogram has been devised for these pigments (Munz and Schwanzara, 1967).

Modulation This refers to a regular change in a stimulus, usually according to a sinusoidal function. Modulation may be spatial, meaning that the light intensity varies as a function of position (e.g. bar pattern or sinusoidal grating), or it may be temporal, meaning that the intensity varies in time (e.g. flickered light). In general, the term modulation is used here to refer to temporal modulation, specifically a mean light level with a superimposed sinusoidal component.

The amplitude of modulation is defined as the amplitude of the sinusoidal component divided by the mean level. For example, if the mean light level were 100 units, and the amplitude of the sine were 10 units (which is half the peak to peak excursion of the sinusoid), the light would vary between a level of 90 units and 110 units, and would be said to have a 10% amplitude of modulation. This corresponds to 0.087 log units.

Peg Experiment A two spot experiment in which the intensity in one of the areas is always held to the same value. See "Experimental Paradigms" in Materials and Methods.

Photopic Refers to the state of light adaptation. In many animals, there is a shift in the spectral sensitivity during dark adaptation; this is usually related to the presence in the retina of rods and cones. Cones are the main contributors to vision in a light adapted condition and for bright lights; they determine the spectral sensitivity in conditions referred to as photopic. In dark adaptation, the spectral sensitivity is determined by the rods, and the condition is termed scotopic. Since there is only one effective photopigment (rhodopsin) at scotopic levels, there can be no color vision; colors are seen only in the photopic condition. In these experiments, the stimuli were of long enough wavelength (710 nm) that they were effective only for the cones containing the long-wavelength pigment; in this sense, all of the work described in this thesis is for photopic stimuli.

Relative Timing An experiment in which a short test flash is presented at various times relative to a longer conditioning stimulus. See "Experimental Paradigms" in Materials and Methods.

Response Segments A form of analysis of the two spot experiments in which responses are observed for particular short intervals during the stimulus, rather than averaging for the entire time of the stimulus. See "Response Segments" in Experimental Results and Conclusions.

Response-Summation Plot A system of plotting the results of the two spot experiments to reveal non-linearities after the first interaction of the two areas. The coordinates are defined as

$$\text{abscissa} = R_1 + R_2$$

and

$$\text{ordinate} = R_{1+2}$$

or, the physiological sum vs the arithmetic sum of responses. See "Method of Response-Summation."

Ricco's Law A psychophysical "law" stating that there is a reciprocal relationship between area and intensity for a threshold response; this holds true only for a limited range of areas. As it is applied to single unit recording, spots of different diameter are placed concentric with the receptive field of the ganglion cell, and the intensity found for a constant (criterion) response. For the range over which the law applies, the relationship is then:

$$A \cdot I = k$$

where A is area, I is intensity, and k is a constant.

Sensitivity The term sensitivity is always used in the strict sense: the inverse of the energy required for a criterion response. For convenience, sensitivity is usually normalized relative to the maximum light available (relative sensitivity); it is also usual to express it as a log relative sensitivity.

As an example, consider a situation in which sensitivity under one condition differs from sensitivity in another such that the stimulus intensity required to elicit the criterion response is not the same in the two cases, i.e.:

$$I_2 = kI_1$$

where I is intensity. The sensitivity, S, in 2 will be 1/k times that in 1:

$$S_2 = 1/I_2, \quad S_1 = 1/I_1$$

$$S_2/S_1 = I_1/I_2 = 1/k.$$

Then the log sensitivity difference will be

$$\log(S_2) - \log(S_1) = -\log(k)$$

Also, note that

$$-\log(I_2) + \log(I_1) = -\log(k)$$

so on a logarithmic intensity coordinate, there will be a separation of log (k) (at the criterion response) between the curves relating intensity of stimulus and magnitude of response (intensity-response curves) for the two conditions. If the two curves are parallel, i.e. may be superimposed by sliding one of them to the left, sensitivity will not be a function of the criterion used. Sensitivity difference is the separation between two such curves.

Sensitivity differences may be due to a number of things. A different steady background illumination could be present, or the physiological state of the retina could be changed. A sensitivity difference is often observed by stimulating two different areas within the receptive field. A plot of sensitivity as a function of position in the receptive field may be drawn; this is called a sensitivity profile.

Sensitivity-Summation Plot A system of plotting the results of the two spot experiments; the summed inverses of the intensities required for a criterion response in each spot are plotted as a function of the intensity required for the same response when both are illuminated. The coordinates are defined as:

$$\text{abscissa} = \log I_{1+2}$$

$$\text{ordinate} = -\log\left(\frac{1}{I_1} + \frac{1}{I_2}\right)$$

or, sum of individual sensitivity of each vs sensitivity of both together. This method reveals non-linearities before the final summation of the two areas. See "Method of Sensitivity Summation."

Shit Term often used in physiology to describe untoward occurrences and results. It particularly applies to situations in which cells are mysteriously lost before they may be recorded, or in which a definitive experiment is completed before it is noticed that the inputs to the tape recorder were disconnected. In its strictest sense, it applies to computer errors. See also "Bloody Hell!"

Shunting A form of inhibitory interaction in which the inhibiting area acts by dividing the signal from the inhibited area.

Spontaneous Rate Most cells continuously fire action potentials in the absence of stimuli. Such maintained activity is known as spontaneous firing. The spontaneous rates vary from unit to unit, but for any given unit, the difference from the spontaneous rate may be considered as a measure of response.

S-Potentials These are slow, apparently intracellular, potentials recorded in the distal part of the retina. When they were first observed by Svaetichin (1953), they were believed to be responses of single receptors; more recent work has shown that this is not so (MacNichol and Svaetichin, 1958), and that S-potentials may be traced to horizontal cells (Kaneko, 1971). S-potentials in goldfish may be monophasic, responding with hyperpolarization for all wavelengths of light (L-type); or they may be color coded (C-type) either biphasic or triphasic (Spekreijse and Norton, 1970). C-type S-potentials respond with either hyperpolarization or depolarization, depending on the wavelength of the stimulus. Biphasic units are hyperpolarizing for one half of the spectrum, depolarizing for the other; triphasic units are depolarizing for green lights, hyperpolarizing for red or blue.

Trichromacy Three independent parameters are necessary and sufficient to describe the normal human color sense. To achieve this, there must be three independent channels of different (but overlapping) spectral sensitivities; the simplest form of trichromacy is one in which there are three separate photopigments. Two important characteristics of trichromatic vision are: 1) all spectral lights are saturated; that

is, no monochromatic light has the appearance of white, and 2) any colored light can be matched exactly by some appropriate mixture of three "primary" lights; the two fields so matched will be indistinguishable, even though their spectral compositions may be quite different. Such a match is called a metameric match.

Transfer Function In general, a transfer function gives the frequency characteristics of a system. Technically, it is the LaPlace transform of the impulse response (see Appendix IV), and may refer to either spatial or temporal frequencies. In the work discussed in this thesis, it is the temporal characteristics that are of interest; the transfer function is the modulation transfer function, or MTF, and is generally determined by measuring the amplitude of the sinusoidal part of the system's response to sinusoidal inputs of various frequencies.

Two Spot Experiment This refers to the experiments in which two separate areas within the receptive field of the ganglion cell were stimulated, either separately or simultaneously. The two particular experiments that fall into this class are the equal-intensity experiment and the peg experiment.

Unit A unit is a cell, usually the cell from which spikes are being recorded.

Table of Symbols Used in Equations

A Area, usually of a circular stimulus

B Intercept of the curve on a sensitivity-summation plot with the ordinate

β Fixed output of a group of receptors in a peg experiment

c Coefficient representing magnitude of subtractive interactions

$\delta(t)$ Dirac delta: unit area pulse of infinitesimal duration

e Base of the natural logarithm system (2.71828...)

e(I) Excitation of a group of receptors

$F(\zeta)$ Non-linear function after summation of excitations from two areas

F(t) Time-dependent multiplier on responses to a step change in intensity

f(s) LaPlace transform of F(t)

f Instantaneous value of F(t)

g Constant, usually the scalar in $F(\zeta)$

γ Exponent in $F(\zeta)$ when it has the form of a power function

I Intensity of illumination

i Subscript variable, refers to the unit under consideration (also, in Appendix IV, the imaginary number $\sqrt{-1}$)

j Subscript variable, refers to other units than the one under consideration

k Constant, usually the constant in the denominator of $F(\zeta)$ when that function is a compressive function

n Exponent of a power function at the receptors

ω frequency, radians/sec

r Radius; distance from center of receptive field

R Response; ganglion cell output

ρ Multiplier (constant) on the excitation from one group of receptors, reflecting a sensitivity difference between receptors from two separate areas

S Sensitivity

s Sum of excitations from two areas, before non-linear $F(\zeta)$; (also complex frequency, $s = i\omega$)

t Time since the onset of the stimulus

v A constant multiplying excitation from both areas; the scalar on e(I)

x Excitation from a group of receptors: $x = e(I)$

x_o Value of threshold

y Excitation from a group of receptors after interaction with other areas and/or passing through a threshold

ζ Dummy variable used as argument of functions

All logarithms are base 10 unless otherwise indicated.

APPENDIX II: SOME ASPECTS OF THE RESPONSE-SUMMATION COORDINATES

Let us consider the simplest model for the center of the receptive field: two spots are illuminated, the receptors at each effect a transformation from light intensity to neural excitation, e(I), the resultant excitations are summed linearly, and a final transformation converts the sum into a spike firing rate. Assume that the transformation at the receptors in each of the areas is the same, and that a sensitivity difference between the two areas may be expressed by a multiplicative constant on the output of one of the areas:

$$e_1(I) = e(I)$$

$$e_2(I) = \rho e(I)$$

(If the form of e(I) is a power function, this representation must be valid; if e(I) is not a power function, the equation above is equivalent to assuming that the sensitivity difference is not due to an intrinsic difference between the cones in the two areas, but to the relative effectiveness of their outputs at the summation point; see footnote, p. 22).

> We are interested in the form of the response-summation plot for two different experiments: equal-intensity and peg (see "Experimental Paradigms").

In the equal-intensity experiment, $I_1 = I_2$; in the peg experiment, I_2 is a constant. Whatever the form of e(I), we may represent the output of the cones from any given area by a quantity x:

$$e(I_i) = x_i \qquad (1)$$

In the absence of a sensitivity difference between the areas of interest, the condition for the equal-intensity experiment is then:

$$x_1 = x_2 = x; \qquad (2)$$

however, in the general case the net excitations from the two areas will be x and $\rho \cdot x$ respectively. For the peg experiment, x_2 is a constant,

and the excitation for that area may be represented as a simple constant: $\rho x = \beta$.

The response to the simultaneous stimulation of both areas is given as:

$$R_{1+2} = F(x_1 + \rho x_2) \tag{3}$$

where F is the transformation after summation (the box marked F in Fig. 4). From equation (2) the ordinate for the curve resulting from the equal-intensity experiment is thus:

$$R_{1+2} = F[(1+\rho)x] \tag{4}$$

and the abscissa is:

$$R_1 + R_2 = F(x) + F(\rho x) \tag{5}$$

Similarly, for the peg experiment, the ordinate is:

$$R_{1+2} = F(x+\beta) \tag{4a}$$

and the abscissa:

$$R_1 + R_2 = F(x) + F(\beta). \tag{5a}$$

Notice that if there were a multiplicative scaler on e(I), its effect would be to change the value of x:

$$vx = ve(I)$$

and therefore also β:

$$v\beta = v\rho x.$$

This is equivalent to multiplying ζ by v in each $F(\zeta)$ above (equations (4), (5), (4a) and (5a)). Such a change of argument affects the range of the function, but does not change the actual curve. If x has a particular range, changing v would cause the plot to trace different segments of the identical curve.

Straight Lines on the Response-Summation Plot

Equal-Intensity Experiment Let us consider what forms of $F(\zeta)$ might yield a straight line for the equal-intensity experiment on response-summation coordinates. For a curve to be a straight line, its derivative must be a constant; that is:

$$\text{slope} = \frac{dR_{1+2}}{d(R_1+R_2)} = \frac{dR_{1+2}}{dx} \cdot \frac{1}{\frac{d(R_1+R_2)}{dx}} = k \qquad (6)$$

where k is a constant. In order to evaluate the numerator and denominator of (6), let us assume that $F(\zeta)$ may be approximated by a series of powers:

$$F(\zeta) = a \cdot \zeta^{n_1} + b \cdot \zeta^{n_2} + \ldots \qquad (7)$$

Substituting equations (4), (5) and (7) into equation (6) gives:

$$\frac{\frac{d}{dx}\left[a(1+\rho)^{n_1} \cdot x^{n_1} + b \cdot (1+\rho)^{n_2} x^{n_2} + \ldots\right]}{\frac{d}{dx}\left[ax^{n_1} + bx^{n_2} + \ldots + a\rho^{n_1} \cdot x^{n_1} + b\rho^{n_2} x^{n_2} + \ldots\right]} = k$$

or, differentiating and collecting terms:

$$\frac{[a \cdot (1+\rho)^{n_1} \cdot n_1]x^{n_1-1} + [b(1+\rho)^{n_2} \cdot n_2]x^{n_2-1} + \ldots}{[a \cdot (1+\rho^{n_1})n_1]x^{n_1-1} + [b(1+\rho^{n_2})n_2]x^{n_2-1} + \ldots} = k \qquad (8)$$

Equation (8) will be valid for all x only if a single power of x appears in both the numerator and the denominator; e.g. if:

$$\frac{(1+\rho)^{\alpha+1} \cdot x^{\alpha}}{(1+\rho^{\alpha+1}) \cdot x^{\alpha}} = k; \qquad (8a)$$

which will be true only if:

$$\frac{dF(\zeta)}{d\zeta} = h \cdot \zeta^{\alpha} \qquad (7a)$$

where h and α are arbitrary constants. Integrating,

$$F(\zeta) = \frac{h \cdot \zeta^{\alpha+1}}{\alpha+1} + k' \qquad \text{if } (\alpha \neq -1)$$

or, for arbitrary g and γ:

$$f(\zeta) = g \cdot \zeta^{\gamma} + k' \qquad (\gamma \neq 0). \tag{9}$$

(Notice that

$$F(\zeta) = h \cdot \log_e(\zeta) + k' \quad \text{if } (\alpha = -1).$$

Thus, a logarithmic function will also give a straight line on the response-summation plot, and will be considered a special case of the power function.)

EQUATION (9) IS THE MOST GENERAL FORM FOR WHICH THE RESPONSE-SUMMATION PLOT WILL BE A STRAIGHT LINE.

The constant k' must be 0 if we are measuring response from spontaneous*, or there would be a response with no input. If k' = 0, equation (4) becomes:

$$R_{1+2} = g(1+\rho)^{\gamma}x^{\gamma} \tag{10}$$

and equation (5) gives:

$$R_1 + R_2 = g(1+\rho^{\gamma})x^{\gamma} \tag{11}$$

or

$$x^{\gamma} = (R_1+R_2)/g(1+\rho^{\gamma})$$

---

*For the purpose of this Appendix, I treat the hypothetical units being modelled as if there were no spontaneous firing. This may be effected in real units by considering response as a difference from spontaneous rate, which is a simple shift of coordinates:

$$R^*_{1+2} = R_{1+2} - d$$

$$(R_1+R_2)^* = (R_1+R_2) - 2d$$

where d is the spontaneous rate. The new coordinates are the ones indicated on each response-summation plot as the "spontaneous" rate lines; the origin of this system is the zero response point.

Substituting into (10):

$$R_{1+2} = \frac{(1+\rho)^{\gamma}}{(1+\rho^{\gamma})} \cdot (R_1+R_2) \tag{12}$$

which is the equation of a straight line through the origin with a slope given by

$$\text{slope} = \frac{(1+\rho)^{\gamma}}{(1+\rho^{\gamma})} \tag{13}$$

Note that equation (13) is independent of x and therefore neither the shape of the curve nor its slope depends on the form of the function at the receptors.

FOR A SYSTEM IN WHICH THERE IS A NON-LINEARITY OF THE FORM OF A POWER FUNCTION AFTER LINEAR SUMMATION, THE RESPONSE-SUMMATION PLOT FOR AN EQUAL-INTENSITY EXPERIMENT WILL BE A STRAIGHT LINE THROUGH THE ORIGIN WITH A SLOPE AS GIVEN BY EQUATION (13).

If $\gamma = 1$, which means the function after summation is linear ($F(\zeta) = g \cdot \zeta$), equation (13) becomes:

$$\text{slope} = \frac{1+\rho}{1+\rho} = 1 \tag{13a}$$

FOR A SYSTEM WITH ALL NON-LINEARITIES BEFORE THE FIRST INTERACTION, THE RESPONSE-SUMMATION PLOT WILL BE A STRAIGHT LINE OF UNITY SLOPE PASSING THROUGH THE ORIGIN.

Peg Experiment Let us now consider the peg experiment, for which

$$\rho x = \beta$$

From equations (4a) and (9) (with $k' = 0$):

$$R_{1+2} = g(\beta+x)^{\gamma} \tag{14}$$

and from (5a) and (9):

$$R_1+R_2 = g(\beta^{\gamma} + x^{\gamma}) \tag{15}$$

then

$$x = \left[\frac{R_1+R_2}{g} - \beta^{\gamma}\right]^{1/\gamma} \tag{16}$$

so, substituting in equation (14)

$$R_{1+2} = g[\beta+\{(\frac{R_1 + R_2}{g})-\beta^{\gamma}\}^{1/\gamma}]^{\gamma} \qquad (17)$$

If $\gamma = 1$, equation (17) reduces to

$$R_{1+2} = R_1+R_2$$

which is the equation of a straight line of unity slope. Notice that this result, like that of equation (13a), is independent of both x and ρ. FOR A SYSTEM WITH ALL NON-LINEARITIES BEFORE THE FIRST INTERACTION OF SEPARATE AREAS, THE RESPONSE-SUMMATION PLOT FOR THE PEG EXPERIMENT WILL BE A STRAIGHT LINE OF UNITY SLOPE PASSING THROUGH THE ORIGIN.

If $\gamma \neq 1$, equation (17) must be used to describe the peg experiment. As an example, let us consider the case for which $F(\zeta) = \sqrt{\zeta}$, or $\gamma = 1/2$. Equation (17) becomes:

$$(R_{1+2})^2 = g[\{(\frac{R_1+R_2}{g}) - \sqrt{\beta}\}^2 + \beta]$$

or

$$(R_{1+2})^2 = \frac{1}{g}(R_1+R_2)^2 - 2 \cdot \sqrt{\beta}\ (R_1+R_2) + 2\beta g$$

which is the equation of a hyperbola, concave upward.
FOR A SYSTEM WITH A NON-LINEARITY OF THE FORM OF A POWER FUNCTION AFTER LINEAR SUMMATION, THE RESPONSE-SUMMATION PLOT FOR THE PEG EXPERIMENT WILL NOT BE A STRAIGHT LINE.

Compressive Function After Summation

Let us now consider the effects of non-linear functions after summation which are not power functions.

ANY FUNCTION OTHER THAN A POWER FUNCTION (OR LOGARITHM) AFTER LINEAR SUMMATION WILL NOT GIVE A STRAIGHT LINE FOR THE EQUAL-INTENSITY EXPERIMENT ON RESPONSE-SUMMATION COORDINATES. The particular function I would like to consider is what I refer to as a compressive function, one that

is linear at low levels, but asymptotes to a ceiling at high levels. The particular form of equation is given in equation (18):

$$F(\zeta) = \frac{g \cdot \zeta}{k + \zeta} \tag{18}$$

where g and k are constants. This is the simplest form of the class of compressive functions that includes the modified power function used by Boynton and Whitten (1970) to describe the receptor function:

$$F(\zeta) = \frac{g \cdot \zeta^n}{k + \zeta^n} \tag{18a}$$

Notice that a function of the form given in (18) operating after a power function gives equation (18a) exactly.

Note also that two functions of the form given in equation (18) acting serially are equivalent to a single compressive function of the same form. If

$$y = \frac{g_1 x}{k_1 + x}$$

and

$$z = \frac{(g_2 y)}{k_2 + y}$$

then

$$z = \frac{[\frac{g_1 g_2}{g_1 + k_2}]x}{[\frac{k_1 k_2}{g_1 + k_2}] + k} \tag{19}$$

If we substitute equation (18) into equations (4) and (5) we find:

$$R_{1+2} = \frac{g(1+\rho)x}{k + (1+\rho)x} \tag{20}$$

and

$$R_1 + R_2 = \frac{gx}{k+x} + \frac{g\rho x}{k+\rho x} \tag{21}$$

To examine the response-summation plot in the vicinity of the origin, we assume that $x \ll k$; equations (20) and (21) become

$$R_{1+2} \simeq \frac{g(1+\rho)}{k} \cdot x \tag{22}$$

and

$$R_1 + R_2 \simeq \frac{g}{k} \cdot x + \frac{g\rho}{k} \cdot x = \frac{g(1+\rho)}{k} \cdot x \qquad (23)$$

Equations (22) and (23) are identical, so

$$R_{1+2} = R_1 + R_2 \quad (x<<k) \qquad (23a)$$

which is a straight line of unity slope through the origin, independent of ρ and x; near the origin, the form of receptor function has no effect on the curve. At high levels, the curve tends to a single terminal point. If x>>k, equations (20) and (21) become:

$$R_{1+2} \simeq \frac{g(1+\rho)x}{(1+\rho)x} = g \qquad (24)$$

and

$$R_1+R_2 \simeq \frac{gx}{x} + \frac{g\rho x}{\rho x} = 2g \qquad (25)$$

So the locus of the final point the curve approaches is one for which the value of R_{1+2} is one half the value of (R_1+R_2). (Note that the actual curve may never reach this point, for x covers a finite range.)

The slope of the response-summation curve is given by equation (6). From equation (20) the numerator of (6) becomes:

$$\frac{dR_{1+2}}{dx} = \frac{kg(1+\rho)}{k^2 + 2k(1+\rho)x + (1+\rho)^2x^2} \qquad (26)$$

and from (21) the denominator of (6) becomes:

$$\frac{d(R_1+R_2)}{dx} = \frac{kg}{k^2+2kx + x^2} + \frac{k\rho g}{k^2+2\rho kx + \rho^2x^2} \qquad (27)$$

As x→0, the slope goes to:

$$\text{slope} \simeq \frac{\frac{kg(1+\rho)}{k^2}}{\frac{kg}{k^2} + \frac{k\rho g}{k^2}} = 1$$

as expected from equations (22) and (23).

If x is very large relative to k, equations (26) and (27) become:

$$\frac{dR_{1+2}}{dx} \simeq \frac{gk}{(1+\rho)x^2}$$

and

$$\frac{d(R_1+R_2)}{dx} \simeq kg\ [\frac{1}{x^2} + \frac{1}{\rho x^2}]$$

so

$$\text{slope} \simeq \frac{1}{(1+\rho)\ (1+\frac{1}{\rho})} \qquad (28)$$

Equation (28) is invariant under replacement of ρ by $1/\rho$; it has a maximum value at $\rho = 1$, slope = 1/4. (Recall that ρ is an arbitrary scaler indicating differences in sensitivity between areas; $\rho = 1$ means the areas are equisensitive). As ρ differs from 1, the slope of the final portion of the curve becomes shallower; the change in slope from the initial portion of the curve is more severe when sensitivities are mismatched.

> Let us now consider the precise shape of the curve on response-summation coordinates when the two areas are equisensitive.

We have seen that ρ has an effect on the shape, but the equations with ρ are unwieldy; since the effect of ρ is only on the sharpness of curvature, let us deal only with the case for $\rho = 1$. In that case, equations (20) and (21) become:

$$R_{1+2} = \frac{2gx}{k + 2x} \qquad (29)$$

and

$$R_1 + R_2 = \frac{2gx}{k + x} \qquad (30)$$

From (30):

$$(R_1 + R_2) \cdot (k + x) = 2gx$$

$$x = \frac{k(R_1 + R_2)}{2g - (R_1 + R_2)}$$

Substituting into (29):

$$R_{1+2} = \frac{2g(R_1 + R_2)}{2g + (R_1 + R_2)} \tag{31}$$

Equation (31) describes the form of the response-summation plot for the equal-intensity experiment on equisensitive areas. Notice that this equation is of the same form as equation (18); also, notice that it is independent of k. Furthermore, x does not appear in (31); no assumption need be made about the form or scale of the function relating light intensity to receptor output. x affects the range of the curve on response-summation coordinates, but not the shape; for different forms of e(I), a different portion of the identical curve would be followed.

The only parameter in the model that appears in equation (31) is g. We have noted that the response to a uniform stimulus changes in time (Figs. 6, 18); if the time-dependent multiplier that causes this effect also causes a change in the shape of the response-summation plot, it must act as a multiplier on g, which is equivalent to multiplying the entire function. If we assume responses to the same stimulus change by a factor f that multiplies g, then from equation (30);

$$R_1 + R_2 = \frac{2fgx}{k + x} = f \cdot (R_1 + R_2) \tag{32}$$

And from equation (29):

$$R_{1+2} = \frac{2fgx}{k + 2x} = f \cdot R_{1+2} \tag{33}$$

Equations (32) and (33) express a simple change of scale on the response-summation coordinates, a dilation or contraction of the plot.

IF THE NON-LINEARITY AFTER LINEAR SUMMATION IS A COMPRESSIVE FUNCTION OF THE FORM

$$F(\zeta) = \frac{g\zeta}{k + \zeta}$$

THE RESPONSE-SUMMATION PLOT FOR THE EQUAL-INTENSITY EXPERIMENT WILL BE A NEGATIVELY ACCELERATED CURVE. AT LOW LEVELS, IT WILL TEND TO A STRAIGHT LINE OF UNITY SLOPE THROUGH THE ORIGIN; AT HIGH LEVELS IT WILL TEND TO A

LIMITING POINT. IF THE AREAS WHOSE RESPONSES ARE BEING SUMMED ARE NOT EQUISENSITIVE, THE SLOPE OF THE CURVE AS IT APPROACHES THE FINAL POINT WILL BE SHALLOWER THAN IF THEY ARE EQUALLY SENSITIVE; IN ALL CASES, THE FINAL POINT TO WHICH THE CURVES TEND IS DETERMINED ONLY BY THE VALUE OF THE SCALER g.

I have not discussed the response-summation plot expected from the peg experiment when there is a compressive function such as the one expressed by equation (18), for the particular result depends on the value of β. One can see intuitively that in the range for which the compressive function is nearly linear, the plots from the peg experiment must coincide with those for the equal-intensity experiment; where the function is quite non-linear, they will not coincide. This is illustrated by the peg experiments shown for the lower two models in Fig. 10.

Lateral Interactions

There are a wide variety of lateral interactions that one could postulate; it clearly would be beyond the scope of this work to demonstrate the features of all the possible forms of interactions that one might consider. The one feature that is shared by all forms of interaction that are non-linear and not mathematically equivalent to a non-linear function after summation is that they will result in a diminution (or enhancement in the case of facilitation) of the responses from each area when both areas are simultaneously active. The extent of the effect may be quite different for different response levels, but the general statement should still apply. Let us therefore consider only one simple case of lateral interactions, subtractive interactions that have been made non-linear by the inclusion of thresholds ($x_o = 0$) in each pathway (see Fig. 4).

It would also be a tedious exercise to consider coupling this interaction with all possible forms of non-linearity after summation.

I shall therefore consider only one function after summation, a power function. This function has the advantage of being relatively simple to analyze, since the resulting response-summation plot for the equal-intensity experiment is a straight line (see above).

Equations (4) and (5) apply in this case, except we must consider the variable y rather than x, where y is the input, after any interactions, to the summation point. When each spot is illuminated alone there is no interaction (because of the thresholds, the lateral influence from any one area stimulated alone is not apparent at the summation point): $y = x$, and equation (5) may be rewritten directly:

$$R_1 + R_2 = F(y) + F(\rho y) = F(x) + F(\rho x) \tag{34}$$

When both spots are illuminated, assuming neither is lit with significantly greater intensity than the other, we may write:

$$\begin{aligned} y_1 &= x_1 - cx_2 \\ y_2 &= x_2 - cx_1 \end{aligned} \tag{35}$$

where c is a constant less than 1 (c is negative for facilitation); these equations assume non-recurrent inhibition.

If we consider the equal-intensity experiment, then $x_2 = \rho x$ and $x_1 = x$. In view of (35), equation (4) becomes:

$$R_{1+2} = F(x + \rho x - cx - \rho cx) = F[(1-c)(1+\rho)x] \tag{36}$$

If $F(\zeta)$ is as given by equation (9), (36) and (34) become:

$$R_{1+2} = g(1+\rho)^\gamma \cdot (1-c)^\gamma x^\gamma \tag{37}$$

and

$$R_1 + R_2 = g(1+\rho^\gamma)x^\gamma \tag{11}$$

From which:

$$R_{1+2} = \frac{(1-c)^\gamma \cdot (1+\rho)^\gamma}{(1+\rho^\gamma)} \cdot (R_1 + R_2) \tag{38}$$

Equation (38) is the equation of a straight line on the response-summation coordinates; its slope is given as

$$\text{slope} = \frac{(1-c)^{\gamma} \cdot (1+\rho)^{\gamma}}{(1+\rho^{\gamma})}$$

This differs from the slope in the absence of lateral interactions (equation (13)) by a factor of $(1-c)^{\gamma}$. If c is positive (but less than 1) the slope will be decreased; if c is negative (facilitation) the slope will increase. For the case of a linear function after summation, ($\gamma = 1$) the slope will be

$$\text{slope} = (1-c),$$

whereas slope is unity without interactions.

To examine the shape of the response-summation plot for the peg experiment, we let $\rho x = \beta$. Applying equation (4) and proceeding as for the equal-intensity experiment:

$$R_{1+2} = g(1-c)^{\gamma} \cdot (x+\beta)^{\gamma} \qquad (39)$$

The responses of each area alone are unaffected by lateral interactions, so equation (15) applies unchanged. Solving it for x (equation (16)), and substituting in (39) we obtain

$$R_{1+2} = g(1-c)^{\gamma}\left[\left\{\frac{(R_1+R_2)}{g} - \beta^{\gamma}\right\}^{1/\gamma} + \beta\right]^{\gamma} \qquad (40)$$

Equation (40) differs from (17) in the factor of $(1-c)^{\gamma}$ again implying a change in slope. If $\gamma = 1$, equation (40) becomes

$$R_{1+2} = (1-c) \cdot (R_1 + R_2)$$

which is a straight line of slope = (1-c) on the response summation plot. NON-LINEAR INHIBITORY INTERACTIONS WILL RESULT IN A DECREASE OF THE SLOPE OF THE RESPONSE-SUMMATION PLOTS FOR BOTH THE EQUAL-INTENSITY AND PEG EXPERIMENTS.

APPENDIX III: SOME ASPECTS OF THE SENSITIVITY-SUMMATION COORDINATES

The sensitivity-summation coordinates are an extension of methods used by Easter (1968a), but he considered summations only between equi-sensitive areas. Sensitivity-summation enables us to consider summation of areas with different sensitivities (see "Method of Sensitivity-Summation").

> We are considering only the equal-intensity experiments in this appendix--the peg experiment would not be expected to give meaningful results on sensitivity-summation coordinates (see "Method of Sensitivity-Summation"). We shall examine the expected form of the sensitivity-summation plot for various permutations of the proposed model of the center of the receptive field (Fig. 4).

Two Spot Summation

Consider the simplest model: two spots are illuminated separately and together, the receptors at each perform a non-linear transformation, the resultant excitations are summed linearly, and a final transformation converts the sum into a spike firing rate. Let us assume (as Easter) that e(I), the transformation at the receptors, may be described as a power function in I, the intensity:

$$e(I) = v \cdot I^n$$

with v and n constants. A sensitivity difference between the two areas may be represented as a multiplicative constant (ρ) on the function for one of the areas (no matter how the sensitivity difference is effected, this representation will be true for a power function; see footnote, p. 22).

$$e_1(I) = vI^n$$

$$e_2(I) = \rho vI^n$$

> The condition for any point on sensitivity summation coordinates is that the responses to each spot alone and both together must be equal;

that is:

$$F[e_1(I_{1+2}) + e_2(I_{1+2})] = F[e_1(I_1)] = F[e_2(I_2)]. \qquad (1)$$

F is a function representing all transformations after the summation, I_1 and I_2 are the intensities in each area illuminated separately, and I_{1+2} is the intensity in each area when both are illuminated simultaneously; I_1, I_2, and I_{1+2} are such that each elicits the same response. Intensity is expressed as quanta$\cdot$sec$^{-1}\cdot$cm^{-2}, so if I_{1+2} were numerically equal to I_1, the retina would be receiving twice as much energy from I_{1+2}; area 1 would receive the identical stimulation in each case, but I_{1+2} also stimulates area 2.

Because we confine the analysis to values of I which give the same response, the inputs to the summation must always be the same. Hence, the transformation F is irrelevant, and we can simplify to:

$$e_1(I_{1+2}) + e_2(I_{1+2}) = e_1(I_1) = e_2(I_2)$$

Substituting the assumed functions e_1 and e_2 we obtain:

$$vI_{1+2}^{\,n} + \rho vI_{1+2}^{\,n} = vI_1^{\,n} = v\rho I_2^{\,n}$$

in which v cancels; there is no effect of scaling both functions, which is in fact equivalent to scaling after the final summation. So:

$$(1 + \rho)I_{1+2}^{\,n} = I_1^{\,n} = \rho I_2^{\,n} \qquad (2)$$

We can now examine the relationship between the intensities required in 1, 2, and 1+2 for the same responses. From (2) it is possible to express I_2 and I_{1+2} in terms of I_1:

$$I_2 = I_1 \cdot (\rho)^{-1/n} \qquad (3)$$

$$I_{1+2} = I_1 \cdot (1 + \rho)^{-1/n} \qquad (4)$$

Let us now examine the coordinates of the sensitivity summation graph. Sensitivity is defined as the reciprocal of the intensity necessary for a given response.

The abscissa is

$$\log I_{1+2} = \log[I_1 \cdot (1 + \rho)^{-1/n}]$$

$$= \log I_1 - \frac{1}{n} \log (1 + \rho) \qquad (5)$$

The ordinate is

$$-\log[\frac{1}{I_1} + \frac{1}{I_2}] = \log[\frac{I_1 I_2}{I_1 + I_2}]$$

$$= \log[\frac{I_1^2 \cdot \rho^{-1/n}}{I_1 + I_1 \cdot \rho^{-1/n}}]$$

$$= \log I_1 + \log[\frac{\rho^{-1/n}}{1 + \rho^{-1/n}}] \qquad (6)$$

Solving for I_1 in (5) and substituting in (6):

$$-\log[\frac{1}{I_1} + \frac{1}{I_2}] = \log I_{1+2} + \log[\frac{\rho^{-1/n}}{1 + \rho^{-1/n}}] + \frac{1}{n} \log (1 + \rho) \qquad (7)$$

which is a straight line of unity slope, and intercept B, where

$$B = -\frac{1}{n} \log\rho - \log[1 + \rho^{-1/n}] + \frac{1}{n} \log (1 + \rho) \qquad (8)$$

Thus the intercept is a function only of the exponent (n) of the receptor function and of any sensitivity difference (ρ) between the two areas. Equation (8) may be rewritten as:

$$B = \frac{1}{n} \log[\rho^{-1/2} + \rho^{1/2}] - \log[\rho^{1/2n} + \rho^{-1/2n}]$$

which has the property of being invariant under replacement of ρ by $1/\rho$. This property is expected from the arbitrary manner in which ρ is defined. Returning to (8), let us examine the simple case of $\rho = 1$--equisensitive spots (such as Easter used):

$$B = -\frac{1}{n} \log(1) - \log[1 + 1^{-1/n}] + \frac{1}{n} \log(1 + 1)$$

$$B = \frac{1}{n} \log 2 - \log 2$$

so

$$n = \frac{\log 2}{B + \log 2} \qquad (9)$$

(This is essentially the relationship used by Easter to estimate the receptor exponent (Equation (6), p. 21). The additional log 2 in the denominator of (9) is a result of a difference between the coordinate system used here and that of Easter: Easter's ordinate was the intensity of a single spot, this ordinate is the negative logarithm of the sum of two sensitivities.

FOR THE CASE OF A POWER FUNCTION AT THE RECEPTORS AND TWO AREAS WHICH ARE EQUALLY SENSITIVE, THE SENSITIVITY-SUMMATION PLOT WILL BE A STRAIGHT LINE OF UNITY SLOPE AND INTERCEPT DIRECTLY RELATED TO THE EXPONENT IN THE RECEPTOR FUNCTION (EQUATION 9).

If the spots are not equisensitive, however, the intercept will be a function of both the receptor exponent, n, and the sensitivity difference, denoted by ρ. What will be the effect on the intercept as ρ takes values other than unity? To find out, let us differentiate (8) with respect to ρ:

$$\frac{dB}{d\rho} = \frac{-\log(e)}{n\rho} - \frac{\log(e)}{(1 + \rho^{-1/n})} \cdot \left[-\frac{1}{n}\right]\rho^{-[(n+1)/n]} + \frac{1}{n} \cdot \frac{\log(e)}{(1 + \rho)} \qquad (10)$$

At $\rho = 1$, $\left(\frac{dB}{d\rho}\right) = 0$, implying a maximum or minimum, as expected from the invariance noted above. To find out which it is, we take the second derivative:

$$\frac{d^2B}{d\rho^2} = \frac{d}{d\rho}\left[\frac{dB}{d\rho}\right] = \frac{\log(e)}{n\rho^2} + \log(e)\left[\frac{\rho^{-(n+1)/n}}{n(1+\rho^{-1/n})}\right]^2 - \frac{\log(e)}{1+\rho^{-1/n}} \cdot \frac{n+1}{n^2} \cdot \rho^{-(2n+1)/n}$$

$$- \frac{\log(e)}{n(1+\rho)^2}$$

At $\rho = 1$,

$$\frac{d^2B}{d\rho^2} = \log(e) \cdot \frac{(n-1)}{4n^2} \qquad < 0 \qquad (0 < n < 1)$$

A negative second derivative implies a maximum at $\rho = 1$; the largest possible intercept is obtained when the spots are equisensitive. If the sensitivities are mismatched, B becomes smaller, resulting in an overestimation of the receptor exponent when equation (9) is used; the receptors will appear more linear than they actually are.
FOR THE CASE OF A POWER FUNCTION AT THE RECEPTORS AND TWO AREAS WHICH ARE NOT EQUALLY SENSITIVE THE SENSITIVITY-SUMMATION PLOT WILL BE A STRAIGHT LINE OF UNITY SLOPE; THE INTERCEPT WILL BE AFFECTED BY THE MISMATCH SUCH THAT THE EXPONENT WILL BE OVERESTIMATED.

Estimation of Receptor Exponent, n

> If two retinal areas under consideration are not equally sensitive, the exponent in the receptor function cannot be accurately estimated from the sensitivity-summation plot alone. An alternative method will be discussed in this section.

If one wants an accurate estimate of n, it could be obtained from equation (8) and the value of ρ, which could be measured directly from the intensity-response curves of the two areas. Alternatively, ρ may be circumvented by a change of coordinates.

Let us rewrite the expression for the ordinate (equation (6)):

$$-\log\left[\frac{1}{I_1} + \frac{1}{I_2}\right] = \log\left[\frac{I_1^{-1} + I_2^{-1}}{2}\right]^{-1} - \log 2 \qquad (11)$$

The expression $\left[\frac{I_1^{-1} + I_2^{-1}}{2}\right]^{-1}$ is merely an average, in this case the harmonic average of I_1 and I_2, or, the inverse of the average sensitivity. (If $I_1 = I_2$, it is obvious that this ordinate is the same as Easter's log I_s, except for the subtracted log 2.) This particular average is merely one member of a class of averages which may be written as

$$I_{av} = \left[\sum_{i=1}^{n} I_i^m/N\right]^{1/m} \qquad (12)$$

in this case, m = -1, N = 2.*

*This form has two advantages over the other obvious case, m = 1, or,

The average to use if one wished to eliminate the dependence on ρ would be to let $m = -n$. In this case the ordinate would be

$$-\log\left[\frac{I_1^{-n} + I_2^{-n}}{2}\right]^{1/n} - \log 2$$

Substituting I_2 from equation (3) in the above expression:

$$\text{ordinate} = -\log[I_1^{-1}] - \log 2 - \log\left[\frac{1+\rho}{2}\right]^{1/n}$$

Since $-\log[I_1^{-1}] = \log I_1$, we can substitute the value of $\log I_1$ from equation (5) so that:

$$\text{ordinate} = \log I_{1+2} + \frac{1}{n}\log(1+\rho) - \log 2 - \log\left[\frac{(1+\rho)^{1/n}}{2^{1/n}}\right].$$

This is the equation of a straight line with unity slope. The intercept, B, is:

$$B = \frac{1}{n}\log(1+\rho) - \log 2 - \log\left[\frac{(1+\rho)^{1/n}}{2^{1/n}}\right]$$

or: $$B = \frac{1}{n}\log 2 - \log 2$$

so: $$n = \frac{\log 2}{B + \log 2} \qquad (9)$$

giving equation (9), independent of ρ.

---

* footnote continued:

expressed as a coordinate:

$$\text{ordinate} = \log [I_1 + I_2] \qquad (13)$$

Equation (11), the ordinate actually used, is a sensitivity average, hence it weights toward the more sensitive area; (13) favors the less sensitive area. The other significant difference is in the direction of errors caused by a sensitivity mismatch. As discussed above, a sensitivity mismatch applied to equation (11) results in an overestimation of the receptor exponent; the receptors seem more linear than they are, so an actual non-linearity might be masked. In equation (13), on the other hand, the exponent would be underestimated; a linear system could look non-linear as a result of a sensitivity difference, and a non-existent effect reported.

This procedure presupposes a knowledge of n, but this is what is being sought. Hence, we must proceed in an iterative fashion, as follows: assume a value of n; solve equation (9) for a better estimate and substitute the newer estimate of n in m = -n; then solve equation (9) again. Starting with an initial assumption of linearity (n = 1), and thereby beginning with sensitivity-summation coordinates (equation (11)), this procedure converges to an estimate of n in two or three iterations.

Lateral Interactions

> The preceding discussion considered only the simplest form of the model; let us now consider the possibility of lateral interactions between area 1 and area 2 before they are summed (see Fig. 4).

As was discussed under "A Model of Retinal Processing," above, interactions which are entirely linear are equivalent to a multiplication on each pathway (v, discussed at the beginning of this appendix), and therefore will have no effect whatever on the sensitivity-summation plots. However, a linear interaction becomes non-linear if a threshold is introduced in each pathway; such an interaction will be considered as a first approximation to any non-linear interaction. The thresholds serve to insure that the lateral interactions will be effective at the summation only when both areas are simultaneously illuminated. Since all they must do is block any negative signals, we shall assume that the threshold value is 0; that is, $x_o = 0$. Larger values of x_o will be discussed later. The interactions are assumed to be symmetrical; each area has subtracted from it the same proportion (c) of the other area's excitation. Notice that if c were greater than 1, there would never be any response to an equal-intensity experiment, for the negative term would always be the larger. If $c < 0$, the interactions are facilitatory.

The conditions for constant response can be expressed as in equation (1), but with the addition of interaction terms when both areas are stimulated together. As above we shall assume e(I) is a power function, and, for simplicity, say the two areas are equisensitive:

$\rho = 1$; because $e_1(I) = e_2(I)$, $I_1 = I_2$.

So:

$$F[e_1(I_{1+2}) - c \cdot e_2(I_{1+2}) + e_2(I_{1+2}) - c \cdot e_1(I_{1+2})] = F[e_1(I_1)]$$

$$= F[e_2(I_2)]$$

or, $$F[2(1-c)e(I_{1+2})] = F[e(I_1)] = F[e(I_2)] \qquad (14)$$

Again, the form of $F(\zeta)$ is irrelevant, so (14) becomes:

$$2(1-c){I_{1+2}}^n = {I_1}^n = {I_2}^n \qquad (15)$$

and

$$I_{1+2} = \frac{I_1}{2^{1/n} \cdot (1-c)^{1/n}}$$

Therefore, the abscissa can be written as:

$$\log I_{1+2} = \log I_1 - \log(2^{1/n}) - \log[(1-c)^{1/n}] \qquad (16)$$

and the ordinate:

$$-\log[\frac{1}{I_1} + \frac{1}{I_2}] = \log I_1 - \log 2. \qquad (17)$$

Solving for I_1 in (16) and substituting in (17):

$$-\log[\frac{1}{I_1} + \frac{1}{I_2}] = \log I_{1+2} - \log 2 + \frac{1}{n} \log (1-c) + \frac{1}{n} \log 2 \qquad (18)$$

Which is again a straight line of unity slope and intercept B, where

$$B = (\frac{1}{n} - 1) \cdot \log 2 + \frac{1}{n} \log (1-c). \qquad (19)$$

Notice that if $c = 0$, B takes the value

$$B = \frac{1}{n} \log 2 - \log 2$$

which leads to equation (9). As c takes values larger than 0 (but less than 1) the right hand term in equation (19) takes increasingly larger negative values; that is, the line on sensitivity-summation coordinates is displaced downward, mimicking a larger value of n in equation (9).

It was this apparent linearizing effect of non-linear interaction that Easter invoked to explain his successful replication of Ricco's law (see "Ganglion Cells" in Introduction).

FOR THE CASE OF A POWER FUNCTION AT THE RECEPTORS AND NON-LINEAR INTERACTIONS (OF THE FORM OF SUBTRACTIVE INTERACTIONS MADE NON-LINEAR BY THE PRESENCE OF THRESHOLDS AT THE SUMMATION POINT), THE INTERACTIONS WILL MAKE THE INTERCEPT ON THE SENSITIVITY-SUMMATION COORDINATES BECOME MORE NEGATIVE, BUT WILL NOT AFFECT THE SLOPE.

Thresholds

We have just seen that thresholds in each pathway before summation can make simple subtractive interactions appear non-linear; however, the thresholds can introduce non-linearities in the absence of lateral interactions. Let us now consider such thresholds (see Figure 13). If x_o, the lowest value of e(I) for which there is an output (y), is larger than 0, the threshold will be of importance at low intensities. The threshold has the form:

$$y = e(I) - x_o \qquad \text{if } e(I) > x_o$$

$$y = 0 \qquad \text{if } e(I) \leq x_o$$

Let us again consider a power function at the receptors, and a sensitivity imbalance ρ; we shall consider a ρ which multiplies each e(I) before the threshold. Let us also consider $\rho > 1$ (obviously, the results must be invariant under replacement of ρ by $1/\rho$). The lowest value of intensity ($I_{1_{min}}$) that will still elicit a response from each spot, separately or both together, will be determined by the less sensitive spot, which is area 1 if $\rho > 1$ (we are not concerned that the response to both at once is due only to one of them, the more sensitive). That is:

$$e(I_{1_{min}}) - x_o = 0$$

or

$$v[I_{1_{min}}]^n = x_o$$

$$I_{1_{min}} = [\frac{x_o}{v}]^{1/n} \tag{21}$$

With thresholds, equation (2) becomes:

$$(1+\rho)I_{1+2}{}^n - \frac{2x_o}{v} = I_1{}^n - \frac{x_o}{v} = \rho I_2{}^n - \frac{x_o}{v} \tag{22}$$

It is now possible, as was done for equations (3) and (4), to express I_{1+2} and I_2 in terms of I_1. The right hand set of relations in (22) is the same as in (2) (x_o/v appears on both sides), so equation (3) still describes I_2 in terms of I_1. However, when I_{1+2} is stated in terms of I_1 we obtain the following result (which is slightly different from equation (4)):

$$I_{1+2} = [\frac{I_1{}^n + x_o/v}{1 + \rho}]^{1/n} \tag{23}$$

The abscissa of the sensitivity-summation plot is therefore:

$$\log I_{1+2} = \frac{1}{n} \log[I_1{}^n + \frac{x_o}{v}] - \frac{1}{n} \log (1+\rho). \tag{24}$$

The ordinate is still as given by equation (6) since the threshold terms appear equally in each of I_1 and I_2 and therefore cancel.

Since equation (24) approximates equation (5) for large values of $I_1{}^n$ relative to $\frac{x_o}{v}$, the plot must look the same as the non-threshold case at high intensities.

The slope of the sensitivity-summation plot is given by

$$\text{slope} = \frac{d\{-\log [\frac{1}{I_1} + \frac{1}{I_2}]\}}{d\{\log I_{1+2}\}} = \frac{\frac{d}{dI_1}\{-\log [\frac{1}{I_1} + \frac{1}{I_2}]\}}{\frac{d}{dI_1}\{\log I_{1+2}\}} \tag{25}$$

which can be obtained by differentiating equations (6) and (24) with respect to I_1. From (6):

$$\frac{d}{dI_1}\{\log[\frac{1}{I_1} + \frac{1}{I_2}]\} = \frac{\log(e)}{I_1} \tag{26}$$

From (24):

$$\frac{d}{dI_1}\{\log I_{1+2}\} = \frac{I_1^{(n-1)} \cdot \log(e)}{I_1^{n} + \frac{x_o}{v}} \tag{27}$$

Finally, the slope is obtained by dividing (26) by (27):

$$\text{slope} = \frac{I_1^{n} + \frac{x_o}{v}}{I_1^{n}} \tag{28}$$

At high values of I_1 the slope is unity. At lower values of I_1, the slope becomes steeper; at the extreme ($I_1 = I_{1_{min}}$), the slope can be obtained from (21) and (28) and is:

$$\text{slope} = \frac{2x_o/v}{x_o/v} = 2$$

The locus of the point on the sensitivity-summation plot corresponding to I_1 at threshold can be found by evaluating the axes at $I_{1_{min}}$. The value of the ordinate from (6) is

$$-\log[\frac{1}{I_1} + \frac{1}{I_2}] = \log(I_{1_{min}}) + \log[\frac{\rho^{-1/n}}{1 + \rho^{-1/n}}]$$

which, from (21),

$$= \frac{1}{n}\log(\frac{x_o}{v}) - \log[1+\rho^{-1/n}] - \frac{1}{n}\log\rho \tag{29}$$

and the abscissa from (24) is:

$$\log I_{1+2} = \frac{1}{n}\log(\frac{x_o}{v}) + \frac{1}{n}\log 2 - \frac{1}{n}\log(1+\rho) \tag{29a}$$

The relationship between the ordinate and the abscissa specifying the locus of the terminal point can be obtained from (29) and (29a):

$$-\log\left[\frac{1}{I_1} + \frac{1}{I_2}\right] = \log(I_{1+2}) + \frac{1}{n}\log\left[\frac{1+\rho}{2}\right] + \log\left[\frac{\rho^{-1/n}}{1+\rho^{-1/n}}\right]. \quad (30)$$

If $\rho = 1$, equation (30) becomes:

$$-\log\left[\frac{1}{I_1} + \frac{1}{I_2}\right] = \log(I_{1+2}) - \log 2.$$

That is, the final point of the curve at $I_{1_{min}}$ can be as far as 0.3 log units below the diagonal. If $\rho >> 1$, equation (30) becomes approximately:

$$-\log\left[\frac{1}{I_1} + \frac{1}{I_2}\right] \simeq \log(I_{1+2}) + \frac{1}{n}\log\frac{\rho}{2} - \frac{1}{n}\log\rho$$

or,

$$-\log\left[\frac{1}{I_1} + \frac{1}{I_2}\right] \simeq \log(I_{1+2}) - \frac{1}{n}\log 2$$

Since $n < 1$ this is still more negative than the previous result. For a square root function ($n=1/2$), the curve could go as far as 0.6 log units below the diagonal if there were a large sensitivity difference; the slope at $I_{1_{min}}$, however, would still be 2.

Notice that there is an alternative explanation for the curve approaching the diagonal at low intensities: the receptor function might become more linear (see "Sensitivity-Summation Plots" in Experimental Results and Conclusions). However, such a change in the receptor function could not account for the curve going below the diagonal.
FOR THE CASE OF A POWER FUNCTION AT THE RECEPTORS, AND IN THE ABSENCE OF LATERAL INTERACTIONS, THE PRESENCE OF THRESHOLDS AT THE SUMMATION POINT AFFECTS ONLY THE LOW INTENSITY PORTION OF THE SENSITIVITY-SUMMATION PLOT. AT HIGH LEVELS OF EXCITATION RELATIVE TO THE THRESHOLDS, THE CURVE WILL BE THE SAME AS IF THERE WERE NO THRESHOLDS (STRAIGHT LINE OF UNITY SLOPE). AT LOW LEVELS OF EXCITATION, THE SLOPE OF THE CURVE INCREASES, AND THE CURVE MAY CROSS THE DIAGONAL OF THE COORDINATE SYSTEM.

Alternatives to the Power Function

Thus far we have dealt only with the case $e(I)=vI^n$; what would be the expected result if e(I) were one of the alternative non-linear functions suggested in Fig. 2?

In the following discussion, we shall only consider the simple case, $\rho=1$, $I_1=I_2$. The first function to consider is a logarithmic function. This function was rejected by Easter on the basis of its predicted behavior on his log I_s vs log I_d coordinates (see "Ganglion Cells" in the Introduction). Assume

$$e(I) = v \cdot \log(kI + 1)$$

(The additive constant (1) is necessary to insure that e(I) = 0 for I = 0; k is a constant.) Equation (2) then takes the form (with $\rho=1$):

$$2 \cdot \log(kI_{1+2}+1) = \log(kI_1 + 1) = \log(kI_2+1) \qquad (31)$$

If we restrict the discussion to intensities such that $kI \gg 1$, (31) becomes:

$$2 \log(kI_{1+2}) = \log kI_1 = \log kI_2$$

then

$$\log I_1 = 2 \cdot \log(I_{1+2}) + \log k \qquad (32)$$

The ordinate of the sensitivity-summation plot is still given by equation (6), and with $\rho=1$ it simplifies to (17). Substituting (32) into (17) gives:

$$-\log[\frac{1}{I_1} + \frac{1}{I_2}] = 2 \cdot \log(I_{1+2}) + \log(\frac{k}{2}) \qquad (33)$$

Thus, for $kI \gg 1$ (that is, $e(I) \simeq v \cdot \log kI$) the expected curve on sensitivity-summation coordinates is a straight line with a slope of 2 and an intercept dependent on the scaling factor, k. Since the observed data did not fall on a line of slope 2, Easter rejected a simple logarithmic function for the receptors. However, a modified form of logarithmic function (such as we are considering here) might be appropriate: at

lower values, the curve will deviate from the prediction of equation (33); as the value of kI approaches 1, the modified logarithmic function closely approximates a power function (see Fig. 2). At these levels, the curve on sensitivity-summation coordinates must approximate a line of unity slope, since that is the curve obtained for a power function at the receptors.

The other non-linear function we shall consider is the modified power function favored by Boynton and Whitten (1970):

$$e(I) = \frac{v \cdot I^n}{k + I^n} \tag{34}$$

At low values of I, this relationship approximates

$$e(I) \simeq \frac{v}{k} \cdot I^n \tag{34a}$$

and must therefore approximate the power function discussed at the beginning of this appendix. At higher levels, however, this approximation fails. Equation (2) then takes the form:

$$2\,\frac{v(I_{1+2})^n}{k+(I_{1+2})^n} = \frac{v \cdot I_1^{\,n}}{k+I_1^{\,n}} = \frac{v \cdot I_2^{\,n}}{k+I_2^{\,n}}$$

so

$$\log I_{1+2} = \frac{1}{n} \log\left[\frac{k\, I_1^{\,n}}{2k+I_1^{\,n}}\right] \tag{35}$$

As before, equation (17) describes the ordinate; differentiating it with respect to I_1 gives the result of equation (26). Differentiating (35) with respect to I_1 gives:

$$\frac{d\log(I_{1+2})}{dI_1} = \frac{2k \cdot I_1^{-1} \cdot \log(e)}{2k + I_1^{\,n}} \tag{36}$$

The slope is obtained from equation (25), with the necessary derivatives given in (26) and (36):

$$\text{slope} = \left\{\frac{2k + I_1^n}{2k}\right\} \qquad (37)$$

For low values of I_1 the slope approaches unity, as we would expect where the receptor function approximates a power function (see equation (34a)). For non-negligible values of I_1, equation (37) becomes greater than unity, the slope increasing with increasing I_1.
FOR THE CASE OF A FUNCTION OTHER THAN A POWER FUNCTION AT THE RECEPTORS, THE SENSITIVITY-SUMMATION PLOT WILL NOT BE A STRAIGHT LINE OF UNITY SLOPE. A PURE LOGARITHMIC FUNCTION WILL YIELD A STRAIGHT LINE OF SLOPE 2; MODIFIED LOGARITHMIC AND POWER FUNCTIONS WILL RESULT IN CURVED LINES ON THE SENSITIVITY-SUMMATION PLOT.

We have now discussed two possible functions that might be postulated for the receptors, and rejected them because of their predicted effects on the slope of the sensitivity-summation plot. From the preceding discussion it might appear that I could now state with mathematical certainty the algebraic form of the early (receptor) non-linearity. From this, the most probable underlying mechanisms could be deduced. Unfortunately, this is not the case in most biological systems, of which this is no exception. We are attempting to select from among a power function (which must certainly fail at the highest intensities, for no retinal unit can produce an infinite output), a modified logarithmic function, and a modified power function. The three are strikingly similar over at least a 2 log unit range (see Fig. 2); they must perforce be quite similar in their effects on the sensitivity-summation plot. The correct inference from the data I have shown is not that the function at the receptors <u>is</u> a square root function, but that it must be a compressive function much <u>like</u> a square root function.

APPENDIX IV: LINEAR SYSTEMS ANALYSIS

Linear systems analysis is a technique of considerable importance in understanding the temporal properties of the visual system* (see Cornsweet, 1970; Creutzfeldt et al., 1970; Grüsser et al., 1970; Maffei et al., 1970; Ratliff, 1965; Spekreijse, 1969; Sperkreijse and Norton, 1970). In this method, the assumption is made that the entire system is linear (or quasi-linear, which means that it behaves linearly over the range of interest), that is:

$$R_{x+y} = R_x + R_y$$

where R_x, R_y, and R_{x+y} are the responses of the system to stimulus x, stimulus y, and stimulus x simultaneous with stimulus y, respectively. It is then obviously true that $R_{2x} = 2R_x$, or in general:

$$R_{\sum_i x_i} = \sum_i R_{x_i} \qquad (1)$$

Since a system's responses are functions of time as well as of the inputs to it, the response at any moment in time to a combined stimulus must be the sum of the system's responses to each alone at that same moment. For this, it is useful to define a property of a system which is called the "impulse response"; this is the system's response (in time) to a unit area pulse: $R_\delta(t)$. Here, $\delta(t)$ is the unit impulse:

$$\delta(t) = 0 \text{ if } \quad t \neq 0$$

and

$$\int_{-\varepsilon}^{\varepsilon} \delta(t)dt = 1, \text{ for all } \varepsilon > 0.$$

For a pulse of amplitude x: $R_x(t) = xR_\delta(t)$.

Any time varying input, x(t), can be expressed as the sum of a series of these impulses. By invoking linearity, as expressed in equation (1),

---

*The technique is also used to analyze spatial properties of the visual system; in such a case, the variables would be expressed in distance along the retina, rather than time.

$$R_x(t) = \int_{t=0}^{\infty} R\delta(\tau)x(t - \tau)d\tau \qquad (2)$$

where τ is the time difference between the occurrence of each $R_\delta(\tau)$ and t.

It is now convenient to transform to the "frequency plane" by applying the LaPlace transform:

$$f(s) = \int_{0}^{\infty} e^{-st} F(t)\, dt \qquad (3)$$

where F(t) is any function in time, and f(s) is the corresponding function in the frequency domain. s is a complex variable; for our purposes, $s = i\omega$, where ω is the frequency of a sinusoidal component, and $i = \sqrt{-1}$. With this restriction, the transform expresses system response in terms of a sum of sines. We can combine equations (2) and (3) to find:

$$r_x(s) = r_\delta(s) \cdot x(s) \qquad (4)$$

By Fourier's theorem, we can expect to be able to find an x(s) corresponding to any input; equation (4) states that the response of a linear system in the frequency domain is the product of its input spectrum and the LaPlace transformation of the impulse response, which is the transfer function. If the system is tested with a simple sinusoidal input there is a single value of ω to be considered, and the amplitude of the sinusoidal output of the system, relative to the input, is the amplitude of the transfer function at that ω. By testing with all possible values of ω, the amplitude and phase characteristics are completely determined; a plot of these characteristics as a function of ω is a Bode plot, also known as the transfer characteristic or modulation transfer function (MTF). The MTF is a complete description of the temporal characteristics of a linear system.

The usefulness of linear systems analysis is confined, naturally, to linear systems, but what has been discussed of intensity-response curves from receptors (see "Receptors") indicates the visual system is non-linear from the very first stage. However, the receptors may be

considered as linear if they are operating over so short a range that the curvature of their intensity-response curves is negligible and may be approximated by a straight line. Practically, it is difficult to achieve a sinusoidal modulation of greater than 90% (that is, the amplitude of the sinusoid is 0.9 times the mean light level). In terms of a logarithmic intensity scale, this corresponds to a peak-to-peak light change of about 1.3 log units. More commonly used modulations have correspondingly lower peak-to-peak excursions: 50% modulation gives about 0.5 log change; 20% yields only 0.176 log. Various functions have been postulated for the receptors; a representative set is shown in Fig. 2. From that figure, it is clear that over these small ranges a straight line is indeed a fair approximation to the curve, and the system may be treated as quasi-linear.

BIBLIOGRAPHY

Abramov, I. (1972). Retinal mechanisms of color vision. In: Handbook of Sensory Physiology Vol. 7/2 Chap 15, M.G.F. Fuortes, ed., Springer-Verlag, Berlin.

Abramov, I., and J. Gordon (1972). Vision and seeing. In: Handbook of Perception Vol. 3 Chap 18 and 19, E.C. Carterette and M. P. Friedman, eds., Academic Press, New York.

Abramov, I., M. W. Levine (1972). The effects of carbon dioxide on the excised goldfish retina. Submitted to Vision Research.

Adams, A. J. (1970). Chromatic, spatial, and temporal influences on single ganglion cell responses of in vivo goldfish retina. (Doctoral Dissertation), Indiana University, Bloomington, Indiana.

Adrian, E. D. and R. Matthews (1927). The action of light on the eye--Part I: the discharge of impulses in the optic nerve and its relation to the electric changes in the retina. J. Physiol. 63, 378-414.

Barlow, R. B. (1967). Inhibitory fields in the Limulus lateral eye. (Doctoral Dissertation), The Rockefeller University, N.Y., N.Y.

Baumgartner, G. (1961). Kontrastlichteffekte an retinalen ganglienzellen: ablietungen vom tractus opticus der katze. In: Neurophysiologie und Psychophysik des Visuellen Systems. R. Jung and H. Kornhuber, eds., Springer-Verlag, Berlin.

Baylor, D. A., M. G. F. Fuortes, and P. M. O'Bryan (1971). Receptive fields of cones in the retina of the turtle. J. Physiol. 214, 265-294.

Boynton, R. M. and D. N. Whitten (1970). Visual adaptation in monkey cones: recordings of late receptor potentials. Science 170, 1423-1426.

Brown, K. T., K. Watanabe, and M. Murakami (1965). The early and late receptor potentials of monkey cones and rods. Cold Spring Harbor Symp. Quant. Biol. 30, 457-482.

Brown, K. T., and K. Watanabe (1965). Neural stage of adaptation between the receptors and inner nuclear layer of monkey retina. Science 148, 1113-1115.

Büttner, U., and O. -J. Grüsser (1968). Quantitative Untersuchungen der raumlichen erregungssummation im rezeptiven feld retinaler neurone der katze. Kybernetik 4, 81-94.

Cavanaugh, G. M. (ed.) (1956). Formulae and methods V of the marine biological laboratory chemical room. Marine Biological Laboratory, Woods Hole, Mass.

Cleland, B. G. and C. Enroth-Cugell (1968). Quantitative aspects of sensitivity and summation in the cat retina. J. Physiol. 198, 17-38.

Cornsweet, T. N. (1970). Visual Perception. Academic Press, New York.

Creutzfeldt, O. D., B. Sakmann, H. Scheich, and A. Korn (1970). Sensitivity distribution and spatial summation within receptive-field center of retinal on-center ganglion cells and transfer function of the retina. J. Neurophysiol. 33, 654-671.

Dartnall, H. J. A. (1953). The interpretation of spectral sensitivity curves. Brit. Med. Bull. 9, 24-30.

Dartnall, H. J. A. (1962). The photobiology of visual processes. In: The Eye, Vol. 2. Part II, H. Davson, ed., Academic Press, N.Y. 321-533.

Daw, N. W. (1967). Color coded units in the goldfish retina. (Doctoral Dissertation), The Johns Hopkins University, Baltimore, Md.

Daw, N. W. (1967). Goldfish retina: organization for simultaneous color contrast. Science 158, 942-944.

Daw, N. W. (1968). Colour-coded ganglion cells in the goldfish retina: extension of their receptive fields by means of new stimuli. J. Physiol. 197, 567-592.

De Testa, A. S. (1966). Morphological studies on the horizontal and amacrine cells of the teleost retina. Vision Res. 6, 51-59.

De Valois, R. L. (1965). Analysis and coding of color vision in the primate visual system. Cold Spr. Hbr. Symp. Quant. Biol. 30, 567-579.

De Valois, R. L., I. Abramov, and G. H. Jacobs (1966). Analysis of response patterns of LGN cells. J. Opt. Soc. Amer. 56, 966-977.

De Valois, R. L., and P. L. Pease (1971). Contours and contrast: responses of monkey lateral geniculate nucleus cells to luminance and color figures. Science 171, 694-696.

Dowling, J. E. (1967). The site of visual adaptation. Science 155, 273-279.

Dowling, J. E. (1970). Organization of vertebrate retinas (the Jonas M. Friedenwald Memorial Lecture). Investig. Ophthal. 9, 655-680.

Dowling, J. E. and B. B. Boycott (1966). Organization of the primate retina: electron microscopy. Proc. Roy. Soc. Ser. B. 166, 80-111.

Dubin, M. W. (1970). The inner plexiform layer of the vertebrate retina: a quantitative and comparative electron microscopic analysis. J. Comp. Neurol. 140, 479-506.

Easter, S. S. (1968a). Excitation in the goldfish retina: evidence for a non-linear intensity code. J. Physiol. 195, 253-271.

Easter, S. S. (1968b). Adaptation in the goldfish retina. J. Physiol. 195, 273-281.

Enroth-Cugell, C. and L. Pinto (1970). Algebraic summation of centre and surround inputs to retinal ganglion cells of the cat. Nature 226, 458-459.

Enroth-Cugell, C. and L. Pinto (1972a). Properties of the surround response mechanism of cat retina ganglion cells and center-surround interactions. J. Physiol. Lond., in press.

Enroth-Cugell, C. and L. Pinto (1972b). Pure central responses from off-center cells and pure surround responses from on-center cells. J. Physiol. Lond., in press.

Enroth-Cugell, C. and J. G. Robson (1966). The contrast sensitivity of retinal ganglion cells of the cat. J. Physiol. 187, 517-552.

Franklyn, J. (1965). Heraldry. A. S. Barnes, Cranbury, New Jersey, 1970.

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

Gordon, J. (1972). Edge accentuation in the frog retina. (In preparation).

Grüsser, O. -J., D. Schaible, and J. Vierkant-Glathe (1970). A quantitative analysis of the spatial summation of excitation within the receptive field centers of retinal neurons. Pflugers Arch. 319, 101-121.

Hartline, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. Am. J. Physiol. 121, 400-415.

Hartline, H. K. (1949). Inhibition of activity of visual receptors by illuminating nearby retinal areas in the Limulus eye. Fed. Proc. 8, 69.

Hartline, H. K. (1969). Visual receptors and retinal interaction. Science 164, 270-278.

Hartline, H. K. and P. R. McDonald (1947). Light and dark adaptation of single photoreceptor elements in the eye of Limulus. J. Cell. and Comp. Physiol. 30, 225-254.

Jacobs, G. H. (1969). Receptive fields in visual systems. Brain Res. 14, 553-573.

Jacobson, M. (1964). Spectral sensitivity of single units in the optic tectum of the goldfish. Quart. J. Exp. Physiol. 49, 384-393.

Kaneko, A. (1970). Physiological and morphological identification of horizontal, bipolar, and amacrine cells in goldfish retina. J. Physiol. 207, 623-633.

Kaneko, A. (1971). Electrical connexions between horizontal cells in the dogfish retina. J. Physiol. 213, 95-105.

Kaneko, A. and H. Hashimoto (1969). Electrophysiological study of single neurons in the inner nuclear layer of the carp retina. Vision Res. 9, 37-55.

Kuffler, S. W. (1952). Neurons in the retina: organization, inhibition and excitation problems. Cold Spr. Hrb. Symp. Quant. Biol. 17, 281-291.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. J. Neurophysiol. 16, 37-68.

Levine, M. W., and I. Abramov (1971). Effects of varying carbon dioxide and oxygen flow on excised goldfish retinae. Read at the Association for Research in Vision and Ophthalmology meeting, Spring, 1971.

Liebman, P. A. and G. Entine (1964). Sensitive low-level microspectrophotometer detection of photosensitive pigments of retinal cones. J. Opt. Soc. Amer. 54, 1451-1459.

MacNichol, E. F. Jr. (1964). Retinal mechanisms of color vision. Vision Res. 4, 119-133.

MacNichol, E. F. Jr. and G. Svaetichin (1958). Electric responses from the isolated retinas of fishes. Am. J. Ophthal. 46, 26-40.

Maffei, L. (1968). Inhibitory and facilitatory spatial interactions in retinal receptive fields. Vision Res. 8, 1187-1194.

Maffei, L. and L. Cervetto (1968). Dynamic interactions in retinal receptive fields. Vision Res. 8, 1299-1303.

Maffei, L., L. Cervetto and A. Fiorentini (1970). Transfer characteristics of excitation and inhibition in cat retinal ganglion cells. J. Neurophysiol. 33, 276-284.

Marks, W. B. (1965). Visual pigments of single goldfish cones. J. Physiol. 178, 14-32.

Marks, W. B., W. H. Dobelle, E. F. MacNichol, Jr. (1964). Visual pigments of single primate cones. Science 143, 1181-1183.

Motokawa, K., Yamashita, E., and T. Ogawa (1960). Studies on receptive fields of single units with colored lights. Tohoku J. Exp. Med. 71, 261-272.

Munz, F. W. and S. A. Schwanzara (1967). A nomogram for retinene$_2$-based visual pigments. Vision Res. 7, 111-120.

Naka, K. I. (1971). Receptive field mechanism in the vertebrate retina. Science 171, 691-693.

Naka, K. I. and W. A. H. Rushton (1966). S-potentials from colour units in the retina of fish (Cyprinidae). J. Physiol. 185, 536-555.

Norton, A. L., H. Spekreijse, M. L. Wolbarsht, and H. G. Wagner (1968). Receptive field organization of the S-potential. Science 160, 1021-1022.

Ratliff, F. (1965). Mach Bands: quantitative studies on neural networks in the retina. Holden-Day, San Francisco.

Rodieck, R. W. and J. Stone (1965). Analysis of receptive fields of cat retinal ganglion cells. J. Neurophysiol. 28, 833-849.

Rosen, P., M. W. Levine, M. Rossetto, and I. Abramov (1970). A system for controlling the light output of a monochromator and for temporally modulating the intensity. Behav. Res. Meth. and Instru. 2, 297-300.

Rushton, W. A. H. and G. Westheimer (1962). The effect upon the rod threshold of bleaching neighboring rods. J. Physiol. 164, 318-329.

Schoenfeld, R. L. (1964). The role of a digital computer as a biological instrument. Ann. N.Y. Acad. Sci. 115, 915-942.

Shapley, R. M. (1971). Fluctuations of the impulse rate in Limulus eccentric cells. J. Gen. Physiol. 57, 539-556.

Snodderly, D. M. (1969). Manufacturing platinum microelectrodes. (unpublished).

Spekreijse, H. (1969). Rectification in the goldfish retina: analysis by sinusoidal and auxiliary stimulation. Vision Res. 9, 1461-1472.

Spekreijse, H. and A. L. Norton (1970). The dynamic characteristics of color-coded S-potentials. J. Gen. Physiol. 56, 1-15.

Spekreijse, H. and T. J. T. P. van den Berg (1971). Interaction between colour and spatial coded processes converging to retinal ganglion cells in goldfish. J. Physiol. 215, 679-692.

Spekreijse, H., H. G. Wagner, and M. L. Wolbarsht (1972). Spectral and spatial coding of ganglion cell responses in goldfish retina. J. Neurophysiol. 35, 73-86.

Sperling, G. (1970). Model of visual adaptation and contrast detection. Perception and Psychophysics 8, 143-157.

Sperling, G. and M. M. Sondhi (1968). Model for visual luminance discrimination and flicker detection. J. Opt. Soc. Amer. 58, 1133-1145.

Stell, W. K. (1967). The structure and relationships of horizontal cells and photoreceptor-bipolar synaptic complexes in goldfish retina. Am. J. Anat. 121, 401-424.

Stone, J. and M. Fabian (1968). Summing properties of the cat's retinal ganglion cell. Vision Res. 8, 1023-1040.

Svaetichin, G. (1953). The cone action potential. Acta Physiol. Scand. 29, Suppl 106, 565-600.

Svaetichin, G. and E. F. MacNichol, Jr. (1958). Retinal mechanisms for chromatic and achromatic vision. Ann. N.Y. Acad. Sci. 74, 385-404.

Thorpe, S. A. (1971). Behavioral measures of spectral sensitivity of the goldfish at different temperatures. Vision Res. 11, 419-433.

Tomita, T. (1970). Electrical activity of vertebrate photoreceptors. Quart. Rev. Biophysics 3, 179-222.

Tomita, T., A. Kaneko, M. Murakami, and E. L. Pautler (1967). Spectral response curves of single cones in the carp. Vision Res. 7, 519-531.

Wagner, H. G., E. F. MacNichol, Jr., and M. L. Wolbarsht (1960). The response properties of single ganglion cells in the goldfish retina. J. Gen. Physiol. 43 part II, 45-61.

Wagner, H. G., E. F. MacNichol, Jr., and M. L. Wolbarsht (1963). Functional basis for "on"-center and "off"-center receptive fields in the retina. J. Opt. Soc. Amer. 53, 66-70.

Walls, G. L. (1942). The Vertebrate Eye. Hafner Publishing Co., New York 1967.

Werblin, F. S. (1971). Adaptation in a vertebrate retina: intracellular recording in Necturus. J. Neurophysiol. 34, 228-241.

Werblin, F. S. (1972). Lateral interactions at inner plexiform layer of vertebrate retina: antagonistic responses to change. Science 175, 1008-1010.

Werblin, F. S. and J. E. Dowling (1969). Organization of the retina of the mudpuppy, Necturus maculosus. II. Intracellular recording. J. Neurophysiol. 30, 546-561.

Wiesel, T. N. and D. H. Hubel (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. J. Neurophysiol. 29, 1115-1156.

Witkovsky, P. (1965). The spectral sensitivity of retinal ganglion cells in the carp. Vision Res. 5, 603-614.

Wolbarsht, M. L. and H. G. Wagner (1963). Glass-insulated platinum micro-electrodes: design and fabrication. In: Medical Electronics, H. Bostem, ed., University of Liege Press, Liege. 510-515.

Yager, D. (1967). Behavioral measures and theoretical analysis of spectral sensitivity and spectral saturation in the goldfish, Carassius auratus. Vision Res. 7, 707-727.

End